

ajer

**THE ALBERTA JOURNAL OF
EDUCATIONAL RESEARCH**

VOLUME XXXV

NUMBER 1

MARCH 1989

**PUBLISHED BY
THE UNIVERSITY OF ALBERTA • EDMONTON**

THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

Published in March, June, September, and December
by the Faculty of Education, University of Alberta

A quarterly journal devoted to the dissemination, criticism, interpretation,
and encouragement of all forms of systematic enquiry into education and
fields related to or associated with education.

EDITOR: Thomas O. Maguire

CONSULTING EDITORS

Reno A. Bosetti
Alberta Department of Education

Robert C. Bryce
University of Regina

Janet A. Donald
McGill University

Michael G. Fullan
University of Toronto

Jane Gaskell
University of British Columbia

Myrna L. Greene
University of Lethbridge

Lilian G. Katz
University of Illinois

Joanne C. Keselman
University of Manitoba

John McLeod
University of Saskatchewan

Alan R. Pence
University of Victoria

Lauran H. Sandals
University of Calgary

Hugh A. Stevenson
University of Western Ontario

Keith C. Sullivan
Dalhousie University

Jaap Tuinman
Simon Fraser University

John D. Wilson
Moray House College of Education

FACULTY ADVISORY COMMITTEE

Harvey Zingle (Chair)

Walter H. Worth

Alton T. Olson

Daiyo Sawada

AJER gratefully acknowledges support from the Social Sciences and
Humanities Research Council of Canada and from the Alberta Advisory
Committee for Educational Studies.

The subscription rate is \$32.00 per year for individuals, \$40.00 per year for
institutions. Add \$8.00 for delivery outside Canada. Single copies are \$8.00
each. Please make cheques payable to *The Alberta Journal of Educational
Research*. Back issues are available; rates supplied on request. Claims for
undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to *The Alberta
Journal of Educational Research*, Faculty of Education, Publication Ser-
vices, 4-116 Education North, University of Alberta, Edmonton, Alberta,
Canada, T6G 2G5.



Faculty of Education
University of Alberta

ISSN 0002-4805

The Alberta Journal of Educational Research

Volume XXXV, Number 1

March 1989

CONTENTS

THE EDITOR'S PAGE: <i>Some Thoughts on Direction</i>	1
Living with Conformity in Student Teaching <i>J.D. MacKinnon</i>	2
Spelling Errors: Grade 1 to University— Analysis of Process and Product	20
<i>P. Evans and A.L. Smith</i>	
The Influence of a Reading Treatment on the Writing Compositions of Grade 11 Students: A Research Update	34
<i>J. Belanger and R.G. Martin</i>	
The Topic Selections and Performance of Adult Student Writers	47
<i>J. O'Shea and K. Beaty Chiste</i>	
Promoting Moral Development Through a High School English Curriculum	61
<i>A. Garrod</i>	
Problem Solving Among Pairs of Achieving and Learning Disabled Students	80
<i>A. Thomas</i>	
Relationships Between Student Perceptions and Teacher Organization in Secondary-Level Alternate School Programs	96
<i>J.B. Woudzia</i>	
BOOK REVIEW	
<i>Education, Research, Information Systems and the North</i> by W.P. Adams	107
<i>Reviewed by R. Carney</i>	

ACKNOWLEDGMENT

The review process is central to the quality of this journal. We thank the following people who reviewed manuscripts in 1988. Their thoughtful critiques contribute to the advancement of theory, research, and practice in education.

Amundsen, C.
Assheton-Smith, M.
Bateson, D.
Berard, R.
Blakey, J.
Brown, D.J.
Burger, J.
Burnett, J.D.
Campbell, C.
Cassidy, J.
Cathcart, W.G.
Chamberlin, C.R.
Clifton, R.
Coleman, P.
Cooper, E.
Crocker, R.
Decore, A.M.
Draper, J.
Evans, P.O.
Ewart, G.
Fagan, W.
Friesen, D.
Gee, T.W.
Gordon, C.
Grant, J.
Greenfield, T.B.
Hague, W.J.
Hayden, R.
Hiebert, B.

Hoge, R.
Housego, B.
Hurlbert, E.L.
Ingram, E.J.
Isherwood, G.B.
Jacknicke, K.G.
Jefferson, A.
Kelsey, G.
Konrad, A.
Kozey, L.
Lang, L.
Larocque, L.
Leithwood, K.
Lucas, B.
MacKay, D.A.
Malicky, G.
Manning, A.
Martin, R.G.
Mattson, N.A.
McLean, L.D.
Miklos, E.
Muir, W.
Mulcahy, R.
Nagy, P.
Neufeld, A.
Olson, A.T.
Peters, F.
Piper, T.
Pollard, M.J.

Randhawa, B.S.
Richards, D.
Ricker, E.W.
Robinson, N.
Romney, D.
Russell, T.
Samiroden, W.
Sawada, D.
Schwarz, A.M.
Stewin, L.
Sullivan, K.
Szetela, W.
Thomas, D.
Titley, B.
Violato, C.
Walker, L.
Webster, J.
Weiss, J.
Wideen, M.
Wilgosh, L.
Willinsley, J.
Wilson, R.J.
Winne, P.
Wright, P.W.
Wright, R.
Young, J.C.
Zatko, G.

THE EDITOR'S PAGE

Some Thoughts on Direction

For the past four years, Wally Worth has guided the fortunes of *AJER*. His legacy is a high quality, well-managed journal that serves the educational research community in Canada and abroad. As the new editor, I intend to continue in the Worth tradition of being a "hands-on" editor. This means that readers will see perspectives from national experts on important research issues, critical reviews of research programs, short notes on promising research directions, and even the occasional long article on a topic which seems fundamental to the development of our field. It also means that writers may experience some frustration as I "encourage" them to make their work more accessible to the wider community of educational researchers.

The main goal of *AJER* is communication. Reports of educational research need not be soporific. Our journal (ours because it belongs to the research community) must stimulate curiosity, excitement, debate, and insight. To further this goal, I am repeating some advice offered by Bill Bruneau, editor of the *Canadian Journal of Education*, who suggested to his contributors that research articles should be clear, should be set in context, should be persuasive, and most importantly, should advance our understanding of education.

Seen from the inside, education is a broad field. Researchers are aware that in a country like ours the number of people who work in a particular area is very small. Few readers of an article will share the specific expertise of the writer. It is important, therefore, to place reports in context: Where does this research fit? Why is it important? Arguments must be carefully crafted and the evidence displayed for the convenience of the reader. The whole point of the report is to provide a coherent, structured portrayal so that readers can assess its significance and decide what relevance it has for their own work.

If we succeed in making our work more readable, and thus more available to scholars working in other areas of education, we will provide a stronger base for the evolution of our field.

T.O.M.

J. DAVID MacKINNON

Acadia University

Living with Conformity in Student Teaching

Despite the criticisms which have been leveled at teacher education programs, the field experience component of preservice preparation continues to enjoy the support of virtually all stakeholder groups. Indeed, it is most commonly identified as an indispensable element of professional preparation. Although graduates are virtually unanimous in their praise of student teaching, questions have been raised about the practicum experience, especially about the tendency of many student teachers to passively conform to the practices of their cooperating teachers. This paper explores the experiences of four prospective teachers of young children during an eight-week practicum, their reasons for conforming, and the implications their stories have for preservice teacher preparation. For these students, conformity was not a matter of passive acceptance of the status quo but was rather a response to their interpretations of the constraints of being a student teacher.

The importance attached to field experiences in the preservice preparation of classroom teachers by all stakeholder groups is a clear testament to the apprenticeship nature of teacher education. Cortis (1985, p. 14) has likened teacher education to medieval craft guilds, and the analogy, at least in its dominant orientation, is not far-fetched. The practicum, specifically, is considered to be the sine qua non of teacher preparation. Lanier and Little (1986, p. 550) point out that while the university-based study of pedagogy is frequently assigned limited value, especially by practitioners, practice teaching is considered to be essential. Its importance is implicitly acknowledged by the sheer quantity of research which has focused on student teaching, while other aspects of professional preparation have received much less attention (Hersom, 1984, p. 255; Nemser, 1983, pp. 155-156). However, the *dynamic* nature of the practicum experience has received scant attention (Adler, 1984, p. 15). As Ryan (1980) has pointed out, "we know very little about the relationship of our training and what is really going on in the minds and emotions of our students" (p. 116).

The purpose of this paper is to explore briefly one aspect of the practicum experience: the oft reported tendency of student teachers to conform to the practices and expectations of their cooperating teachers (Beyer, 1984; Clifton, 1979; Dickie, 1983; Goodman, 1985; Seperson & Joyce, 1973; Tardif, 1984; Zimpher, DeVoss, & Nott, 1980). In essence, it is an attempt to come to some understanding of the dynamics of student teaching. Why is this important? One could cite the relatively dismal knowledge base in teacher education (Ralston, 1980; Sears, 1984), the questionable impact of professional preparation on teaching practice (Arnstine, 1979; Lortie,

J. David MacKinnon is an assistant professor in the School of Education at Acadia University, Wolfville, Nova Scotia.

1975), the suggestion that the practicum is primarily a socialization (some have suggested an uncritical one) into the world of schools (Iannaccone, 1963; Smith, 1980), or the alleged theory-practice dichotomy which finds frequent expression in the teacher education literature (Miklos & Greene, 1987; Miller & Taylor, 1984). All of these are supportable reasons for investigating just what is going on when education students venture into real classrooms with real children.

But there is another reason as well. Tabachnick, Popkewitz, and Zeichner (1979-80) maintain that "the characteristics of the student teaching experiences are not to be found in public statements of intention but through an examination of the experience itself" (p. 28). It is the individual's *interpretation* of his or her experience that is so important in moving toward an understanding of the world through another's eyes, in determining how someone has understood what they have experienced. Teacher education programs, at the very least, are intended to have a beneficial and transforming influence on prospective teachers. Given the stated importance of the practicum and its potential for facilitating (or thwarting) the transfer of theory into practice, teacher educators should never assume that field experiences are in and of themselves worthwhile and educative. Nor should they assume that what is intended to happen during the practicum in fact happens.

Changing orientations to educational research are now opening the doors to a number of studies which have as their fundamental intent an understanding of the dynamics of student teaching (Britzman, 1986; Dickie, 1983; Feiman-Nemser & Buchmann, 1987; Lapin, 1984; Ralston, 1980; Tabachnick et al., 1979-80; Tardif, 1984; Zeichner & Liston, 1987). Some of these studies raise serious questions about the value of student teaching as it is presently structured in many universities. The data presented in this paper are intended to contribute to our understanding of preservice field experiences and their value in preparing prospective teachers.

The Nature of the Study

Introduction

The data on the practicum presented in this article derive from a larger study of the teacher education experience (MacKinnon, 1987). The purpose of this larger study was to contextualize and interpret the meaning which selected teacher candidates gave to their university classroom and practicum experiences during one academic term. In order to do this I spent one academic term with a class of 27 third-year education students specializing in early childhood at a Canadian university.

By the time the practicum began, I had already spent a number of weeks with the class as a participant observer. I had attended all their classes, taken part in all their activities (except compulsory assignments and tests), and spent innumerable hours with a reduced number of them exploring what course work in teacher education was like. (They were enrolled in two courses: one on program environments in early childhood education, and one in educational administration.) Who I was and what I was doing was

well known to all of them; they had welcomed me into the class. Being with them so much of the time allowed me to be present as experience unfolded and they began to talk about, and react to, their day-to-day activities. It also placed me in a unique position to develop the solid field relationships so crucial to the success of a study of this type (Whyte, 1979, p. 59). Consequently, as they prepared to begin student teaching, I had already come to know most of the students reasonably well and had developed rapport with several of them. This helped me greatly when it came time to select individuals who would become key informants during the practicum.

A Conceptual Orientation

As *meaning* was the key concept in the study, I drew on the basic tenets of symbolic interactionism, a major social theory which addresses the interpretation of meaning. Interactionism takes as its central focus the relationship between the individual and the larger social setting. This perspective essentially rests on three underlying assumptions about the individual in society: (a) that people act toward things based on the meaning these things have for them; (b) that the meaning of such things comes from the social interaction one has with others; and (c) that these meanings are handled and modified through an interpretive process (Blumer, 1969, p. 2). This interpretive process, conceptualized as intermediary between stimulus and response, is critical in symbolic communication, for the same symbol can have different meanings in different contexts (Ritzer, 1983, p. 163). Thus response is conceptualized as a constructed act based on the meaning which the individual has assigned through an interpretation of significant symbols.

Among the many concepts native to symbolic interactionism, one that emerged as particularly relevant to the practicum experience, and is thus of central importance here, is "definition of the situation." Definition of the situation refers to the way individuals take into account the various elements in their immediate situations in order to decide on appropriate behavior. As Lauer and Handel (1983) note, this involves representing the situation "to the self symbolically so that a response can be made" (p. 127). The significance of how individuals define situations is underscored in the classic assertion by Thomas and Thomas (1928) that "if men define situations as real, they are real in their consequences" (p. 572). How situations are defined is influenced by a host of factors including the goals one has in a particular situation, the application of perspectives developed from previous interaction, and the view one has of oneself in that setting (Charon, 1979, pp. 134-138).

Situations may also be defined intentionally. That is, individuals pursuing particular goals in given situations will attempt to influence the definitions of others. This may be achieved by what one chooses to say, by how it is said, by certain physical gestures, and even by choice of clothing. This is especially prevalent in Goffman's (1959) dramaturgical approach to interactionism. The drama metaphor is founded on the concept of impression

management, that individuals attempt to guide and influence the impressions others form of them.

For the participants in this study, the ways in which they defined the practicum had a significant bearing on their decisions regarding what they could and could not do as student teachers.

The Fieldwork

This paper focuses exclusively on the practicum. Student teaching occupied the last eight weeks of the academic term, and this particular practicum was especially important in the minds of all the students because they had been placed at their preferred level, that is, K-3. The majority, however, found themselves in kindergarten classrooms, with only a few placed at any of the other three grade levels.

In order to make the study manageable, I focused exclusively on four student teachers. Three of them were in a kindergarten setting, and one in grade 1. Two factors influenced the selection of these students. The first involved the nature of our relationship. I selected individuals with whom I related easily and who indicated that they would be willing to accommodate regular classroom visits from me. The second pertained to the nature of their placement. In early childhood classes at the university, a great deal of emphasis had been placed on the structuring of appropriate learning environments for young children. ("Structure" was an overriding theme throughout the term, and it referred to the way a learner's time is engaged while he or she is in the classroom or learning environment. A highly structured classroom is one in which a student exercises little self-direction and most of the activities are controlled by the teacher.) The principle thrust of these classes had been that relatively unstructured, child-directed activities are more suitable than structured, teacher-directed ones. The four student teachers I chose to visit had been placed in classrooms in which teaching arrangements varied from relatively unstructured to tightly structured (or traditional) techniques. Most of these set-ups would not have been condoned by their early childhood instructors and I wanted to see how the students dealt with their situations in light of this, or whether it had an influence at all.

I visited each student teacher at least once each week for a number of hours, and all conversations were recorded (with permission) on a pocket-size cassette tape recorder. The tape recorder also served other purposes. It was a substitute for my field book (which I had used extensively during university classes), thus allowing me to chronicle events or make personal comments about things that happened. In addition, I would talk with each informant during the evening after an observation in order to explore the day's events more fully. These telephone calls were all recorded.

In order to supplement the data I gathered during my visits and conversations with these four student teachers, I selected eight students to serve as secondary informants. I interviewed each of these student teachers once a week over the telephone and recorded the conversations. (In the case of two of these secondary informants, we met at a prearranged location every

week or more for two to three hours of informal discussion on their student teaching experience.) In two instances I made an on-site visit. The data gathered from these conversations were used to supplement and enrich that collected from the four key informants.

In order to ensure that what I was seeing and understanding of the practicum experience was in fact what the student teachers were experiencing, each was asked to critically review all quotations attributed to them and the context in which they were being employed. The credibility of data of this nature lies in determining whether categories created in data analysis are shared in meaning between the researcher and the participants (LeCompte & Goetz, 1982, p. 44).

Data Analysis

The general orientation which I adopted toward data analysis followed from a conceptualization of research as more a cognitive undertaking and less a validation process. Bailyn (1977, p. 97) points out that this implies data analysis as an ongoing, continuous activity involving a constant interplay between concepts and data in which the conceptual understanding of the researcher is developmental rather than assumed beforehand. Such an orientation allows continuous data analysis to guide the inquiry by moving from broad descriptive observations to more focused and selective ones (Spradley, 1980). For each student teacher, I maintained a chronological record of the events which took place during the eight-week practicum. This record was composed of the numerous tapes made during school visits and evening telephone conversations. Each tape was dated and marked with the name of the student teacher recorded thereon. In addition, each student teacher maintained a log of her activities during the practicum, and I was permitted access to these.

The actual analysis of these data was a task far greater in magnitude than I had ever imagined. Although I had read numerous accounts by other researchers who had undertaken similar projects, their descriptions of the difficulties encountered held little meaning for me until I saw the amount of data that was accumulating.

I attacked this task by first reviewing the suggestions made by Bogdan and Biklen (1982) and Turner (1981). Both offer useful advice for reducing large quantities of written material to a more manageable form. Turner (1981) sets out a nine-stage prescription for the manipulation of raw data into refined analysis which, although not followed precisely, was useful to me in the early coding and categorization of my data. I shied away, however, from rigidly following any prescriptive account of "qualitative" data analysis, for this seemed too much like prestructuring what should not be prestructured.

Initially I planned to have the tapes transcribed. Then I was going to code and analyze the transcripts in a manner similar to the way I had dealt with my field notes during the first half of the term. But transcribing proved to be both expensive and time-consuming. The time element was important, for I needed to listen to all the tapes from one week before I visited

the student teachers again the next week. After a period of trial and error, I settled on the following system: After every visit and follow-up telephone conversation, I listened to the tapes and made notes on the issues and concerns identified by each individual. For each note I made, I recorded its tape footage so that I could easily retrieve the reference. When I finished this I gave each issue or concern which I had noted a descriptive label or category and entered each of these on its own file card. In the end, I defined over 30 categories for each student teacher—each of which represented an issue that she had identified during the practice teaching experience. The whole process was begun anew for each individual in order to avoid the temptation to make the categories developed for one student teacher fit another.

Once this was done I looked for similarities between categories for each individual and, as a result of apparent interrelationships, reduced the numerous categories into a smaller number of themes. One of these themes forms the basis of later discussion on the student teaching experience.

Before presenting the data on the student teaching experience, I wish to provide a brief picture of the four principal participants as they began the practicum. Although the names have been changed, these descriptions are nonetheless intended to give body and context to what would otherwise be nothing more than a collection of accounts from faceless, nameless students.

The Key Informants

Becky

Becky was slightly older than most of her classmates. She had been married and had worked at other jobs before making the decision to become a teacher. Nevertheless, she claimed she had always wanted to be a teacher and fondly recalled playing school with her younger brother when she was a child.

She was placed in a kindergarten classroom in one of the area's newer schools. There were two kindergarten classes each day, one in the morning and one in the afternoon, and each lasted for two and a half hours. Each class was structured in a similar manner. The children spent the first 30 minutes in the gymnasium where they either played on their own or engaged in exercises aimed primarily at gross motor development. From there they went to their classroom for various opening exercises and 15 to 30 minutes of teacher-directed time ("circle time"). Then came approximately 30 minutes of "center time," where they could play on their own at learning centers which had been set up around the room. The session wound down with a short recess and ended with a 30-minute block of teacher-directed time.

Becky was really excited about her placement. She had nothing but praise for the atmosphere in the school and for her cooperating teacher in particular. This is how she described it when I first visited her:

It's really neat. The whole atmosphere in the school ... the kids are free to be people. They're not told to sit down or shut up, as you can tell by this room (referring to the flexible arrangement in the room). It's the way a school should be—fun. It still has

rules, it still has structure, but it's fun. It's just neat to see. I hope that's the way our trend is going.

Beth

Unlike most of the other student teachers, Beth claimed to have made the choice to become a teacher about two months before arriving at university. However, she recalled that as a child she always assumed the role of leader when she and her family visited her younger cousins. Throughout childhood and adolescence she received much positive reinforcement about her ability to relate to young children.

Beth was placed in a kindergarten classroom in an inner-city Catholic elementary-junior high school. There were two kindergarten classes each day. The first 30 minutes of class began with teacher-directed opening exercises including attendance, prayer, a good morning (afternoon) song, calendar activity, and a check on the weather. The next 30 minutes was devoted to center time, when the children either visited learning centers of their own volition or engaged in a more structured activity such as an art project. This was followed by a snack and a 15-minute recess. The day ended with more teacher-directed activities such as music time, story time, and a good-bye song.

Beth's initial reaction to her placement stood in marked contrast to Becky's. Her dismay can be directly traced to the university's early childhood classes she had attended. There the students had been consistently exposed to a child-centered philosophy of school curricula and pedagogy which taught that children should be permitted to actively engage in the learning process and that they are, if given the chance, capable of providing the self-direction to do so. Consequently, "appropriate" early childhood classrooms were ones which allowed for a lot of self-direction by the children and minimized the amount of time where the teacher directly controlled what was taking place.

Beth was undoubtedly the most fervent champion of an open approach to teaching in the class, and she was astounded and disappointed to find herself in a kindergarten where most of the children's time was teacher-directed. She was especially distressed to learn that her cooperating teacher's approach to language instruction centered on individual letters of the alphabet: letter days, cutting out letters, tracing letters, learning the sounds of the letters, and so on. She ardently believed that it was all meaningless. Because of this, the tone set early in the practicum was anything but positive (although she did not openly betray her feelings to her cooperating teacher). This is how Beth put it:

My initial reaction was that (in a slightly subdued voice) this is horrible. Because it's very quiet in here. You know, the kids do not get overly loud and have fun and play.

Margo

Margo, like Becky, was slightly older than most of her classmates. She was a daycare worker before entering university. But like Beth, her decision to become a teacher could not be traced to a childhood desire. In fact, while

teaching was something she could see herself doing, she also had other interests: "If I never taught a day in my life it wouldn't bother me."

She was placed in a kindergarten classroom in a small suburban elementary school. There were, again, two classes each day, and the structuring of the school day was similar to Becky's and Beth's situations. The first 30 minutes was for teacher-directed circle time. The routine was fairly structured: a few moments with the calendar, then on to the weather; a review of hand signals (which told them to look or to listen), and a song or two. This was followed by 30 minutes of center time. The children could go to various centers, but there were always at least two which were compulsory (usually math and language). After a short recess the children went to the library where they could browse on their own or listen to a story. The day concluded with a 30-minute gym period.

Margo's belief in a child-centered approach to curriculum and instruction was much like Beth's. However, she had a happy-go-lucky side to her personality which seemed to prevent her from becoming as distressed as Beth over the negative aspects of her placement. She quickly adjusted to the fact that, although this was not how she would structure her own classroom, this was the way things were and there was little she could do about it. But she told me she was disappointed at first, and her feelings about her placement were unambiguous in the following log entry, written on the second day of the practicum:

I feel that [this] classroom is geared for the teachers. There is not one speck of children's work on any part of the classroom walls. I feel strongly that the classroom should be the children's room, that anyone who would walk into the classroom would be able to get a feel for the children in the room because the walls would be full of their work (stories, pictures, writing, art, etc.). I find that this room has many lovely posters and pictures on the walls—but do the children notice them or make comments about them? Many of them are not even on the child's eye level. How meaningful are these wall decorations? I know that this is something I would change in my own room.

Melinda

Melinda was one of the younger members of the class and had always wanted to be a teacher. She clearly remembered playing school as a youngster, and while she could not isolate a particular time when she made the decision to become a teacher, she felt she had always been headed in that direction.

Unlike the other three, Melinda was placed in a grade 1 classroom in a Catholic elementary school, and it contrasted with the other placements in almost every way. Her day was much more traditionally arranged. Instead of preparing for a two and a half hour class, Melinda was faced with a school day which was divided into eight periods. While a specialist teacher taught music, and her cooperating teacher retained the daily religion class (Melinda was not Catholic), Melinda, once she assumed a full load, was responsible for language arts, math, science, social studies, health, art, and physical education. Most of these subjects occupied a 40-minute time slot each.

Melinda was the only one who did not react noticeably to the structural elements of her placement. Whereas Margo and Beth were decidedly critical of their situations, and Becky ecstatic with hers, Melinda rarely mentioned the physical arrangement or time structuring of her grade 1 classroom. She was neither critical nor approving. Instead, she was more concerned from the beginning with gaining a sense of what her cooperating teacher expected from her and how she would be evaluated. In her words:

I was just more concerned with what I would be asked to do, [and] when. What was my teacher like? How would she evaluate me? How much flexibility would she allow me in planning my lessons and things like that?

Student Teaching

During the fieldwork and data analysis stages of the research, I developed eight themes on the practicum experience based on what I was seeing and hearing during the hours spent with each student teacher. Each theme represented, by definition, a unifying idea running through a portion of the data. Specifically, each was something shared by all the student teachers, and while they did not necessarily experience it in a similar fashion, it was an issue or concern for all. These eight themes focused on (a) the issue of conformity (the focus of this paper); (b) the significance of the cooperating teacher to the student teaching experience; (c) the tensions created by university supervisors, especially where their expectations for the student teachers were at variance with those of the cooperating teachers; (d) the experience of being evaluated by both cooperating teachers and university supervisors; (e) the importance to the student teachers of having effective classroom control; (f) attempts by some of the student teachers to subtly change some of the structural arrangements in their classrooms; (g) the tendency to view student teaching as a performance; and (h) the distinctions made by many of the student teachers between the “real” world of the classroom and the “ideal” world of the university.

A dominant theme, that is, one that arose again and again in conversation with the student teachers, was their sense of having to conform to the expectations (stated or perceived) of their cooperating teachers. All of them, some more than others, were critical of the pedagogy they saw practiced in their classrooms. Most would have liked to try out some of the approaches advocated in the early childhood classes they took at university, and they claimed that if they were in their own classrooms, things would be different. But they were not in their own classrooms, and they did not try out many new ideas. On the surface, it looked as if the hours spent in classes had had little effect. What follows is a glimpse below the surface. It is an attempt to understand why they conformed and how they felt about doing what they did. And in the telling, their stories raise questions about what purpose a practicum is intended to serve in the preparation of classroom teachers.

Role Definition: A Tendency to Conform

A question of status. Guinea pigs: That was how Margo described student teachers during the first few days of a practicum. She felt that because they were new on the job the children would try to test them, and because

the student teachers did not yet know the rules of the school and the classroom, they had to fumble along as best they could.

From the beginning Becky, Beth, Margo, and Melinda found that their status as student teachers was really not a status at all. The classrooms they entered were not theirs. As such, they were guests, and they saw themselves as being subservient to their cooperating teachers. This meant holding in check all the impulses and beliefs which might clash with what they defined as the existing norms of their situations. Margo told me early in the practicum that she had no choice about doing certain things in her room. I asked why.

(She laughs.) I just wouldn't go and say: "I'm totally against it. I'm not going to do it." You have to realize that I'm a student teacher. I'm not running the classroom. You don't go into the class and start dictating what you like and what you don't like. There's a lot of things I would do differently but that doesn't mean I can change it.

Margo was the best at articulating what it was like to be a student teacher, although each of the others told the same story in different words. It was like being expected to be a teacher and a student at the same time: a teacher to the children, but a student to the cooperating teacher. But, for Margo at least, the student role dominated. As she said: "I think until I'm not a student I'll feel like a student." Near the end of the practicum she wrote in her log: "I still don't feel 100% comfortable at all times. Perhaps because I know that I am playing the role of a *student teacher*." And even when she did feel certain of herself in her opposition to some aspect of classroom practice, she usually chose to comply rather than make an issue of it. "You just kind of learn to bite your tongue or play along. It's not worth the time and effort to make a big deal out of things that I don't feel comfortable with."

She wasn't alone in this understanding of what it was to be a student teacher. Melinda shared her views, and although it snowballed to the point of causing her a great deal of irritation, she felt powerless to do anything about it:

[Sometimes] I feel frustrated and restricted and ... well, not overwhelmed ... just getting tired of people telling me what to do. Just having people above you all the time, you know, it's not like it's your own classroom. Sure I'm going to be [student] teaching full time, but it's not my classroom. There's rules that have been established in there that I don't feel I have any right to change.

Beth also felt this sense of powerlessness in the way she defined the role of student teachers. And, given her dedication to unstructured learning environments, this was a bitter pill to swallow. Her classroom was, in her view, anything but unstructured. Nevertheless, she felt she had no option:

You've got to do what the teacher says. You're stuck. You can't be your own teacher; you've got to do what the teacher says so that you're not looking like you're coming in here and saying: "Well, your program sucks. I'm doing it my way."

So compliance was, in the eyes of these student teachers, a taken-for-granted part of being an outsider in someone else's classroom.

Many of the secondary informants concurred. One told me how frustrated she was at having to conduct lessons which focused on individual let-

ters of the alphabet (a no-no according to their early childhood instructors) and frequently referred to them as “stupid.” But she didn’t voice her complaints to her cooperating teacher: “I just do what she tells me to do and everything’s OK.” And when I would question this approach, she would tell me that that was what her teacher had planned for the year, and that’s the way it was. It was as if there were no other possible answer, as if my question were silly. Another said of her situation, referring to her cooperating teacher, “It is her class.” And yet another: “It’s her classroom. I’m just the student teacher. I’m there to learn from her.” Everyone said the same thing.

The impact of change. But there was more to conformity for student teachers than simply defining their status as being subservient to their cooperating teachers. The impact of any potential change on the children was also on the minds of many of them. I first became aware of it in conversation with Becky:

I like to pay attention to how she [the cooperating teacher] deals with the kids. After you choose the teacher, she’s the boss. The way she does it is the right way as far as these little kids are concerned. It’s the only way they’ve ever seen. So you basically have to model your teaching in accordance with this. If it’s your classroom and you set up the routine, the rules and the way you teach with the kids, they’re used to you. So you have to respect that and you can’t go doing things differently. So I like to watch and listen to the way she talks with the kids and pick up on how she deals with them because they are comfortable with that, and I don’t want to make them feel uncomfortable.

At first I wondered whether this was little more than a way of rationalizing compliance with established procedure. But all of them told me of instances where children were quick to speak out when their student teacher inadvertently strayed from the standard routine. On some occasions I observed this myself. The children in Melinda’s grade 1 classroom told *her* how she was supposed to mark their assigned work, and Becky’s kindergarten children informed *her* how “show and tell” was to be conducted. Beth told of mistakenly changing the pattern of morning exercises one day by starting the morning song before saying the prayer. The children interrupted her to say that things were not done that way. She tried to save face by asking if it was all right to change the pattern for just this one day. Most went along, but even then some adamantly began reciting the prayer during the song. She did not deviate again.

In addition, Beth was concerned that change would lessen the children’s self-esteem. She told of doing a phonic routine with the class where she showed the letter of the day to the children and they made the sound. It concerned her that they frequently made the sound before the letter was visible. She thought of changing the order but was fearful that the children would think that what they thought they knew yesterday they did not know today. So the impact of change was a concern to her.

Evaluation. The most frequently mentioned reason for conforming to established practice, however, was the fact that all of them were being evaluated by their cooperating teachers. Long before the practicum began, each of the student teachers “knew” that it would be the make or break

component of their four-year BEd program. This had been well ingrained through discussions among peers, reports from others who had gone before, advice from experienced teachers, and comments by their university instructors. In fact, once their placements had been posted, one of their early childhood instructors announced in class: "Your student teaching is what they look for in job placements." This was not news to anybody.

Beth was the most ardent champion of the early childhood "philosophy." There was no doubt in her mind that learning experiences had to be structured in a meaningful way for young children. To her, this meant unstructuring them, that is, allowing the children to assume a significant amount of control over what they did during the school day. Knowing this, I asked her why she abided by the structural arrangements present in her classroom:

Because they're evaluating you. If they don't like you they're not going to give you a good evaluation no matter how good you are. That's the sad truth but I think that is the truth. You talk to people who've gone through it before and people say: "Well, I've had a conflict of interest with my teacher." And it shows on their final evaluation.

Margo agreed:

Having people observe, or knowing you're being evaluated holds me back, I know it does. Because I know if I did it I would do things the way *I* felt right and there would be no one questioning me. I think that makes a difference. Until I'm not a student, I feel like a student.

In her eighth week, Margo's frustration was plainly evident: "*Let me teach one whole week just for myself.* I don't want anyone observing."

Becky's situation was somewhat different. Her classroom came closer to being an unstructured setting than any of them. Furthermore, she had developed a warm and friendly relationship with her cooperating teacher. So on the rare occasions when she found herself opposed to some detail of pedagogy, she chose not to confront her teacher. But she made it abundantly clear that even if she was opposed to *most* of what her cooperating teacher was doing, she would play along. A good evaluation was an important part of the student teaching experience.

Evaluation was mentioned frequently by many of the secondary informants. I met with two of these student teachers one weekend for a three-hour conversation about their experiences. At one point during our discussion, one of them referred sarcastically to a lesson she had conducted on the letter N: "I did it exactly as she [her cooperating teacher] did it." Her friend, who was seated across the table from her, made a symbolic gesture which, she explained, meant "brown-noser." The other responded: "Listen, I want to get a good mark. I don't have any choice."

The Effect of Conforming

Conformity did not mean an abandonment of their own beliefs about appropriate teaching styles. Beth referred to student teaching as a "bending to mold yourself into the mold that the teacher has." But it had an impact. As the practicum wore on, Margo found herself becoming more and more resigned to doing just as her cooperating teacher asked without even so

much as a murmur of discontent. During my last visit in her eighth week she told me of the details of a particular instructional unit she was doing. She said that her cooperating teacher had suggested a number of ideas that she might incorporate into the unit. Margo commented, almost smugly, "Well, you know what I'm going to do? Exactly that! Because I know that's what she expects."

Not surprisingly, Beth had the greatest feelings of constraint regarding her situation. She was decidedly unhappy with what she saw happening to herself as the weeks passed. During the fifth week I asked her if she had experienced any conflict situations since my last visit. She replied:

The usual. Doing things I don't like doing, like doing the letter *R*. I was thinking this morning that I came in just like a robot and changed the letter of the week, and got out the stuff and thought about how I was going to do the letter *R*. And it didn't even strike me as: "God, this is horrible," until after a while I thought: "What am I doing?" I'm not even thinking about it anymore. I'm into this rut where I'm doing it, and it scares me.

Conformity was, simply put, a fact of life for the student teachers throughout the eight-week practicum. Whether for reasons of status, or out of concern for the children, or as a result of a pragmatic self-interest in a good evaluation, all of the student teachers defined the practicum as a situation where significant change was not an advisable course of action. This meant following the established schedule, maintaining the existing structure, and even, in some instances, adopting certain mannerisms of their cooperating teachers. Put another way, student teaching simply did not provide the opportunity for these prospective teachers to try out many of the ideas and skills they had learned at university. None of them would argue that the experience was a waste of time (although Beth came close), but all of them viewed the practicum as an artificial teaching experience in some ways. With the possible exception of Becky (who had an exemplary relationship with her cooperating teacher), none felt that she could truly be herself as long as she was constantly being scrutinized and asked to perform in a setting that was not hers. There was some evidence of overt contradiction, usually slight structural alterations by the student teachers in favor of a more child-centered approach to learning and teaching. But for the most part, these changes took the form of seemingly innocuous deletions or insertions in an existing routine. Perhaps Margo was not being too harsh when she described student teachers as guinea pigs.

Discussion

As I mentioned at the beginning of this paper, conformity is not an unknown aspect of student teaching. In a study of 10 student teachers at a large American university, Goodman (1985) found that the majority of them "passively" accepted the traditional curriculum and standard forms of instruction. He does refer to one particular individual, whom he describes as "intellectually alive and socially aware," who sought out a cooperating teacher who would let her experiment during practice teaching (p. 45). But even here the student felt compelled to adhere to certain aspects of the prescribed curriculum.

Beyer (1984) also comments on the tendency of student teachers to abide by existing instructional arrangements, linking it to a desire on behalf of the students to succeed as professionals. But he worries that this desire, coupled with a cooperating teacher who reinforces compliance with the status quo, may lead to “uncritical uniformity” (pp. 36-37).

And in a study of four second-language student teachers, Tardif (1984) noted their tendency to conform to standard practice in their classrooms by referring to it as “taking the path of least resistance” (p. 152). It is clear, however, that they disapproved of some of what they did and did it only because to do otherwise might jeopardize success during the practicum. She quotes one of her informants:

All the cooperating teachers I had were very different and I think I did well with all of them. The reason I did well was because I made a point to adapt myself to them and their situation. I think you have to do that as a student teacher. If you're not prepared to do that then be prepared to have a tough time. (p. 152)

This was essentially what Beth was saying when she spoke of “molding yourself” to the cooperating teacher’s ways of doing things. But the students in my study did not take the “path of least resistance” simply because it was the easiest way to go. Nor did they “passively” accept the traditional curriculum and standard forms of instruction, although if one were to judge solely on the basis of overt behavior, it would appear that way. And there was no “uncritical uniformity,” although Becky’s situation approached it, but I think this was so because her classroom was less structured than the others and thus more in line with what she considered to be an appropriate learning environment. On the contrary, most of these student teachers, and especially Beth and Margo, were highly critical of what they witnessed during the practicum. But the significance of the practicum to their teaching careers, at least as they defined it, all but precluded any attempts to do anything other than what the cooperating teacher specified. For them, the practicum was not an opportunity to try out techniques they had spent hours discussing and on which they had prepared assignments and presentations at university. It was, instead, a time for observation and replication.

Rethinking Student Teaching

It has been two years since I conducted the fieldwork for this study. Three of the four early childhood students are now practicing teachers. Becky and Beth are teaching kindergarten, and Melinda is teaching grade 1. Margo has not decided on a definite career move yet. She is an adventurer, and at the time of this writing is doing just that—having adventures—in a foreign land. But when she returns, assuming a teaching position is a real possibility.

I do not know what impact their student teaching experiences will have on them as full-time teachers. If their comments and criticisms from the practicum are accurate indicators of future actions, many will attempt to incorporate into their instructional repertoires much of the child-centered, child-directed orientation advocated in university classes and to minimize the more structured activities they noted during the practicum. Yet some

researchers have pointed out that the “liberalized” attitudes of education students during their preservice years is little more than a temporary adoption of university ideology which is quickly replaced by deeply-ingrained traditional perspectives once that experience ends (Lortie, 1975; Shipman, 1967). Nonetheless, Zeichner and Tabachnick (1981, p. 10) contend that enough questions have been raised about this scenario to warrant some skepticism about the supposed lack of impact of university coursework.

At the beginning of this paper I pointed out that field experiences occupy a prominent place in the preservice training of teachers. Even though teacher education is currently coming under regular attack, particularly in the United States, most critics agree that whatever else might be dispensable, practica are not (Cruickshank & Armaline, 1986, p. 35). But what is their purpose? If it is simply to acquaint prospective teachers with current practice, as Beyer and Zeichner (1982, p. 19) contend is frequently the case, then what is currently being practiced in many universities is probably adequate. Ratzlaff and Grimmett (1985, p. 27), however, infer that if this is the case, if professional preparation is to be merely an uncritical, passive socialization into the world of teaching, then teacher education might best be removed from academic settings and accomplished through field-based programs at the hands of experienced teachers.

If, on the other hand, practica are meant to reinforce the transfer of theory into practice, then the nature of that practice must come under careful scrutiny, and the experience must be structured to maximize the accomplishment of this goal. In my particular study, the early childhood instructors at the university hoped that the practicum would provide an opportunity for their students to try out much of what they had learned. This did not happen, and so regardless of whatever else may have been accomplished by this process, the fundamental purpose of the field experience was undermined.

There are, of course, other things that student teaching does. Cruickshank and Armaline (1986, p. 36) list many of them including providing an opportunity for education students to receive feedback on their performance, adding realism and meaning to preservice preparation, offering an opportunity for novices to gauge whether they have selected the appropriate career, and affording an occasion for students to learn the “elements of a profession.” But all of this could be accomplished by some form of apprenticeship program that has no connection whatsoever with a university. The richness of the university experience comes from repeated exposure to theoretical and conceptual issues, and the *raison d’être* of an undergraduate school of education within a university is to apply current theoretical thinking to educational concerns and to demonstrate their application to practice. If any component of a teacher education program erodes this basic function, it needs to be either restructured or eliminated.

There is a problem, however, if we end our discussion on this note. Surely a mindless application of instructional strategies learned in university classes would be as counterproductive as an uncritical replication of current practice. Every effort should be made to place student teachers in situations

where they will be encouraged to experiment with new techniques and to explore different ways of approaching problems which arise in class. This means, among other things, that they must have a solid grounding in curricular and pedagogical theory. But the task, as Craib (1984, p. 5) points out in a discussion of social theory, is less one of learning theories and more one of learning to think theoretically. He likens this to "learning a new language in a particularly difficult way, not by building up vocabulary and learning the various grammatical rules, but by listening to the language being spoken, in all its complexity, its slang, dialects, etc." (p. 5). As I see it, this sort of systematic and rigorous approach to understanding social phenomena shares some similarities (although it is not the same thing) with what Morgan (1983, p. 374) refers to as "critical reflection," where the individual (in our case, the student teacher) is encouraged and assisted to explore the theories and assumptions which underlie a particular action, the consequences of that action, and alternative ways of acting. This may be a particularly promising avenue of exploration for making the student teaching experience an educative one. For while critical reflection is unlikely to overcome conforming tendencies among student teachers, encouraging and assisting students to carefully analyze the assumptions which underlie classroom actions and to consider alternative instructional approaches may go a long way toward ensuring that conforming actions as student teachers do not follow them into their own classrooms.

There is evidence that practica are being rethought and restructured along these lines in an attempt to make the student teaching experience an educative one. Zeichner and Liston (1987) provide an excellent account of one program in the United States that builds on the concept of "reflective action." Its essential purpose is to assist prospective teachers in seeing the assumptions that underlie educational practice, and to provide an environment where they are encouraged to do much more than simply replicate existing practice. The student teachers are expected to live within the frameworks already in place in the classrooms where they teach, but they are also expected to explicate basic assumptions about teaching and learning that are inherent in school activities, to develop plans for modifying and going beyond specified materials, to maintain a journal, and to meet for weekly seminars at the university to assess and explore what they are seeing and doing in the schools. The authors make it clear that there are many hurdles to overcome, but the program, if nothing else, is a move in the direction of *guiding* (and this means far more than just *supervising*) the student teaching experience, rather than assuming that a field experience will be beneficial, *ipso facto*.

Beyer (1984) describes another program at an American university that ostensibly shares much in common with the one described by Zeichner and Liston (1987). The emphasis again is on critical reflection as a means of assisting students in seeing beneath the surface of what happens in schools. The benefit of this approach, as the authors point out, is that "the involvement in critical reflection and field experience ... goes some way toward

redirecting the tendencies of utilitarian teaching perspectives, excessive realism, and uncritical replication" (p. 40).

The point of these two examples, and a lesson to be learned from the experiences of the participants in this study, is that the student teaching experience should be as much a part of university-based, university-controlled teacher education as taking a course in early childhood or educational administration. The kinds of activities described by Zeichner and Liston (1987) are an example of this. If field experiences are not clearly defined this way in the minds of all concerned, if they are understood to be in some way divorced from university work, then the notion of transferring current theory into practice becomes little more than empty words, and the practicum opens itself to being a miseducative experience at best.

References

- Adler, S. (1984). A field study of selected student teacher perspectives toward social studies. *Theory and Research in Social Education*, 12(1), 13-30.
- Arnstine, D. (1979). The temporary impact of teacher preparation. *Journal of Teacher Education*, 30(2), 51.
- Bailyn, L. (1977). Research as a cognitive process: Implications for data analysis. *Quality and Quantity*, 11(2), 97-117.
- Beyer, L.E. (1984). Field experience, ideology, and the development of critical reflectivity. *Journal of Teacher Education*, 35(3), 36-41.
- Beyer, L.E., & Zeichner, K.M. (1982). Teacher training and educational foundations: A plea for discontent. *Journal of Teacher Education*, 33(3), 18-23.
- Blumer, H. (1969). *Symbolic interactionism: Perspective and method*. Englewood Cliffs, NJ: Prentice-Hall.
- Bogdan, R.C., & Biklen, S.K. (1982). *Qualitative research for education: An introduction to theory and methods*. Boston: Allyn & Bacon.
- Britzman, D.P. (1986). Cultural myths in the making of a teacher: Biography and social structure in teacher education. *Harvard Educational Review*, 56(4), 442-456.
- Charon, J.M. (1979). *Symbolic interactionism: An introduction, an interpretation, an integration*. Englewood Cliffs, NJ: Prentice-Hall.
- Clifton, R.A. (1979). Practice teaching: Survival in a marginal situation. *Canadian Journal of Education*, 4(3), 60-74.
- Cortis, G. (1985). The closure and re-organisation of institutions of teacher education in England and Wales, 1972-1981. *Journal of Educational Thought*, 19(1), 13-23.
- Craib, I. (1984). *Modern social theory*. Brighton, UK: Wheatsheaf Books.
- Cruickshank, D.R., & Armaline, W.D. (1986). Field experiences in teacher education: Considerations and recommendations. *Journal of Teacher Education*, 37(3), 34-40.
- Dickie, B.T. (1983). *Portrait of a practicum*. Unpublished doctoral dissertation, University of Alberta, Edmonton.
- Feiman-Nemser, S., & Buchmann, M. (1987). When is student teaching teacher education? *Teaching and Teacher Education*, 3(4), 255-273.
- Goffman, E. (1959). *The presentation of self in everyday life*. Garden City, NY: Doubleday Anchor Books.
- Goodman, J. (1985). What students learn from early field experiences: A case study and critical analysis. *Journal of Teacher Education*, 36(6), 42-48.
- Hersom, N. (1984). Teacher education: Research priorities and organizational means. In P. Grimmett (Ed.), *Research in teacher education: Current problems and future prospects in Canada* (pp. 254-256). Vancouver: University of British Columbia.
- Iannaccone, L. (1963). Student teaching: A transitional stage in the making of a teacher. *Theory into Practice*, 2(2), 73-80.
- Lanier, J.E., & Little, J.W. (1986). Research on teacher education. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 527-569). New York: Macmillan.
- Lapin, F.M. (1984). *Professional socialization: A naturalistic study of students becoming teachers*. Unpublished doctoral dissertation, University of Arkansas, Fayetteville.

- Lauer, R.H., & Handel, W.H. (1983). *Social psychology: The theory and application of symbolic interactionism* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- LeCompte, M.D., & Goetz, J.P. (1982). Problems of reliability and validity in ethnographic research. *Review of Educational Research*, 52(1), 31-60.
- Lortie, D.C. (1975). *Schoolteacher: A sociological study*. Chicago: University of Chicago Press.
- MacKinnon, J.D. (1987). *Early childhood teacher education: Contextualizing and interpreting the student experience*. Unpublished doctoral dissertation, University of Alberta, Edmonton.
- Miklos, E., & Greene, M.L. (1987). Assessments by teachers of their preservice preparation programs. *Alberta Journal of Educational Research*, 33(3), 191-205.
- Miller, P.J., & Taylor, G.D. (1984). The theory-practice relationship in teacher education: Need there be any? *Alberta Journal of Educational Research*, 30(4), 332-339.
- Morgan, G. (1983). Toward a more reflective social science. In G. Morgan (Ed.), *Beyond method* (pp. 368-376). Beverly Hills: Sage.
- Nemser, S.F. (1983). Learning to teach. In L.S. Shulman & G. Sykes (Eds.), *Handbook of teaching and policy* (pp. 150-170). New York: Longman.
- Ralston, S.D. (1980). *Learning to teach: An ethnography of student teachers' perspectives*. Unpublished doctoral dissertation, Michigan State University, East Lansing.
- Ratzlaff, H.C., & Grimmett, P.P. (1985). *Role expectations for the three positions in the student teaching triad*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, 1985. (ERIC Document Reproduction Service No. 264 226)
- Ritzer, G. (1983). *Contemporary sociological theory*. New York: Knopf.
- Ryan, K. (1980). Inside the black boxes: The process of teacher education. In G.E. Hall, S.M. Hord, & G. Brown (Eds.), *Exploring issues in teacher education: Questions for future research* (pp. 105-118). Austin: University of Texas at Austin.
- Sears, J.T. (1984). *A critical ethnography of teacher education programs at Indiana University: An inquiry into the perceptions of students and faculty*. Unpublished doctoral dissertation, Indiana University, Bloomington.
- Seperson, M.A., & Joyce, B.R. (1973). Teaching styles of student teachers as related to those of their cooperating teachers. *Educational Research*, 31(2), 146-151.
- Shipman, M.D. (1967). Theory and practice in the education of teachers. *Educational Research*, 9(3), 208-212.
- Smith, B.O. (1980). On the content of teacher education. In G.E. Hall, S.M. Hord, & G. Brown (Eds.), *Exploring issues in teacher education: Questions for future research* (pp. 7-34). Austin: University of Texas at Austin.
- Spradley, J.P. (1980). *Participant observation*. New York: Holt, Rinehart and Winston.
- Tabachnick, B.R., Popkewitz, T.S., & Zeichner, K.M. (1979-80). Teacher education and the professional perspectives of student teachers. *Interchange*, 10(4), 12-29.
- Tardif, C. (1984). *On becoming a teacher: The student teacher's perspective*. Unpublished doctoral dissertation, University of Alberta, Edmonton.
- Thomas, W.I., & Thomas, D.S. (1928). *The child in America*. New York: Knopf.
- Turner, B.A. (1981). Some practical aspects of qualitative data analysis: One way of organising the cognitive processes associated with the generation of grounded theory. *Quality and Quantity*, 15(3), 225-247.
- Whyte, W.F. (1979). On making the most of participant observation. *American Sociologist*, 14(1), 56-66.
- Zeichner, K.M., & Liston, D.P. (1987). Teaching student teachers to reflect. *Harvard Educational Review*, 57(1), 23-48.
- Zeichner, K.M., & Tabachnick, B.R. (1981). Are the effects of university teacher education "washed out" by school experience? *Journal of Teacher Education*, 32(3), 7-11.
- Zimpher, N.L., deVoss, G.G., & Nott, D.L. (1980). A closer look at university student teacher supervision. *Journal of Teacher Education*, 31(4), 11-15.

PETER EVANS

ARDYS L. SMITH

University of Victoria

Spelling Errors: Grade 1 to University— Analysis of Process and Product

Research in spelling is often focused on either spelling products or spelling processes. Likewise, the common view of spelling is dichotomous: a word is either spelled correctly or incorrectly. To assess the relationship of process and product and to identify developmental trends in spelling, the unconventional spellings which occurred in 200 short essays written by students in grades 1, 4, 8, 10, and first-year university were examined and categorized by strategy employed (process) and error type (product). Analyses of strategy and error type by grade level suggested that (a) the categories employed could be used to reliably analyze spelling behavior; (b) there is a notable developmental trend in both strategy and error type; (c) certain orthographic patterns are particularly difficult for learners at various stages; and (d) employment of certain strategies results in specific types of errors. Findings indicate that there is a strong process/product relationship.

Introduction

Historically there have been two major approaches to the investigation of spelling. One approach has centered on the strategies employed by spellers; the other, more common, approach has focused on the error patterns evident in written samples. Although researchers using the process or the product approach have generally mentioned the relationship of process and product, little has been done to systematically study the combination of the two.

Teachers and the public have usually viewed spelling as a simple dichotomous process; the spelling of a word was either correct or incorrect. That is, only the products of attempts at spelling were considered. However, more recent research in spelling has focused on process, that is, the place that spelling holds in the child's overall attempt to learn language. Seen from this perspective, spelling ability grows as the result of the application of certain semiotic principles, as noted below, in the child's attempt to communicate.

The semiotic view proposes that spelling is a single case of a more general process—the use of a symbol system to convey meaning. There are essentially two parts to this view of the acquisition of early spelling competence: development of an awareness of the nature and use of symbols and

Peter Evans, a professor of language arts at the University of Victoria, has published numerous articles on aspects of language learning and has collaborated on writing student texts for the junior high school reading program.

Ardys Smith is a doctoral student in language arts at the University of Victoria.

the subsequent use of articulatory features for selection of symbols. In the first phase children recognize the operation of symbol systems by noticing that symbols are carriers of meaning in the world around them. At this stage, children will use any symbol in their attempts to convey meaning (Harste, Burke, & Woodward, 1983). This interest in the use of symbols soon grows to include an awareness that the graphic symbol system—conventional orthography—is such that certain symbols are conventionally used to represent certain sound values. In the next phase, youngsters begin to attempt to combine their knowledge of the articulation of speech sounds and the names of letters used to represent these sounds in their writing (Read, 1975).

Ehri (1980) describes the sophistication of this knowledge based on examples reported by Read:

The pre-school child is very sophisticated in analyzing a word's phonetic structure and in inventing clever and plausible ways of representing that structure orthographically. In fact [he or she] appears to be better at this than adults who have not been trained in phonetics. The pre-school orthographer's spellings change systematically as [he or she] gains in orthographic spelling. (p. 325)

Spelling is one part of the semiotic process in which the child attempts to convey meaning through the use of conventional symbols. Within this process, spelling competence appears to be developmental; that is, there are distinct stages in the learner's approximations of conventional spelling (Beers & Henderson, 1977; Zutell, 1980). Invented spellings appear to reflect youngsters' categorization of speech sounds based largely on articulatory features of these speech sounds. As well, spelling entails both a *process*, the employment of strategies, and a *product*, the written word.

The objective of this research was to replicate and expand Read's (1975) and Bebout's (1985) studies and to incorporate the semiotic view presented in the work of Harste et al. (1983). Whereas Read and Harste et al. focused on the spelling patterns and strategies of young children, this research was designed to investigate the developmental changes in these patterns from first grade to first-year university. In addition, the results of students' spelling attempts were analyzed to provide data on the products of these spelling strategies. Process and product were then compared.

In summary, the study addressed five basic questions:

1. Are the findings of Read, Harste et al., and Bebout replicable?
2. How do the spelling strategies and spelling products of students change over time?
3. Do these changes provide evidence of developmental trends?
4. Are particular types of words especially difficult for learners at various stages?
5. What are the relationships between spelling processes and spelling products?

Background to the Study

A review of the related literature indicates that spelling has only recently become an area of major interest to researchers in the language arts. This

current interest is largely focused on the process of learning to spell, that is, on youngsters' use of invented spellings and the strategies employed by pre-school and primary-grade children.

Beers and Henderson (1977) and Gentry (1978, 1984) have identified and described the developmental aspects of youngsters' growth in spelling competence. Gentry (1984) compared these stages of oral language acquisition and defined the stages as precommunicative, semiphonetic, phonetic, transitional, and correct.

Frith (1980) found that a majority of the errors produced by 12-year-olds who were poor spellers but good readers were phonetic misspellings. Frith's analysis of these misspellings led to a description of three stages in the process of spelling: correct analysis of speech sounds; conversion of phonemes to graphemes; and selection from all the phonetically plausible graphemes. Frith's description of these stages in the development of abilities in spelling supports the view of the interaction of process and product in the development toward the use of conventional orthography and guided the analysis of spelling in this study.

Read (1975) investigated the phonetic stage of spelling ability and described the categorization of speech sounds by four- to seven-year-olds. He provided evidence for youngsters' use of articulatory features in the categorization and representation of speech sounds in spelling.

A majority of the studies of process and product in children's spelling, including those cited above, have been conducted in experimental settings. Harste et al. (1983) attempted to overcome the limitations of the experimental paradigm by studying youngsters' spelling as it occurred in natural writing situations. This qualitative approach, in which the subjects served as informants, provided additional insight into the strategies employed in the spelling process. Harste et al. identified six spelling strategies: spelling the way it (1) sounds, (2) articulates, (3) looks, or (4) means; (5) spelling in the way similar problems have been solved; and (6) spelling by the child's own or someone else's rules.

There are a number of areas in which the strategies identified by Harste et al. and the patterns described by Read overlap. Both identified strategies related to spelling sounds using similar letter names and spelling based on articulatory features. Read developed a variety of articulatory subcategories, for example, tense/lax vowels and preconsonantal nasals, while Harste et al. grouped all the spelling errors based on articulation into a single category.

In contrast to the work of Gentry, Beers, and Henderson, Read, and Harste et al., which focused on the spelling *process*, Bebout (1985) investigated error types, that is, *products*, generated by learners who spoke English as their primary language and those for whom English was a second language. Although Bebout's primary purpose was to compare the error patterns of the two groups, her findings related to the native English speaking group, thus her categorization schemes were useful in this study.

The eight major categories of errors identified by Bebout are (1) consonant-doubling errors, (2) other consonant errors, (3) errors involving

schwa, (4) errors involving silent *e*, (5) other vowel errors, (6) letter misordering, (7) unclassifiable items, and (8) homophones.

Anderson (1985) provided a detailed overview of recent research into the development of spelling strategies and stated that "Children constantly revise their rule systems.... Their strategies ... change as they gain more control of written language and become more aware of spelling principles and conventions" (p. 141). These strategies were described as changing from attempts to spell based on perceived graphophonic relationships to greater reliance on semantic or structural analogies to the use of lexical and visual information.

The preponderance of evidence in studies conducted over the last decade supports the view of spelling as a developmental process, with similarities of stages, strategies, and patterns across large numbers of children in their attempts to use English orthography to convey meaning. This evidence was used as the basis for seeking developmental patterns in this study.

Design

The Research Model

The research design for this study can be described as a one-shot case study (Campbell & Stanley, 1963). In this design each group is studied only once subsequent to some agent or treatment presumed to have brought about the observed behavior. In the present study the treatment was the subjects' previous learning history in spelling, and the observed behavior was the kind of spelling errors made in the course of writing compositions.

Subjects and Procedures

Short essays were collected from students in each of grades 1, 4, 8, 10, and first-year university. The university samples were acquired from University of Victoria students attending introductory composition classes. First-grade samples were chosen from a large collection of writing done by students in the Victoria School District. The grades 4 through 10 samples were randomly selected from a large number of samples collected throughout the province as part of the British Columbia Writing Assessment. These essays were designed to assess the writing competence of students and taxed them to do their best work in terms of content, style, and mechanics including spelling. Students were encouraged to use the most effective words possible to express their ideas; consequently, they appeared willing to use words they could not spell correctly. This approach was assumed to provide the researchers with the broadest possible sample of spelling attempts. One purpose of this study was to assess spelling strategies employed in actual writing activities rather than in artificially constructed activities in which the spelling of specific words is targeted. Fifty samples of writing from each grade level, for a total of 200 samples, were analyzed.

All spelling errors from each sample were listed and categorized for inclusion in selected categories established by Read (1975), Harste et al. (1983), and Bebout (1985). Comparisons were made across spelling strategy, spelling strategy by grade level, and between spelling strategy and error type. All analysis of spelling errors into error categories and strategies

was done by the research assistant, a doctoral student, and the principle investigator. Interrater reliability was established at .91.

Determination of the strategies used was done by examining and categorizing the types of errors found in the misspelled words on the assumption that these categories reflect the students' approaches to orthographic representation. In other words, if students appeared to make an error based on the sound of the word in their attempts to spell it, it was assumed that the use of graphophonic relationships was one of their spelling strategies. The strategy categories identified for examination were taken from the work of Harste et al. (1983), Gentry (1978, 1984), and Read (1975) and include spelling a word by (1) how it sounds, (2) how it is articulated, (3) its appearance, (4) its meaning, (5) analogy with other spellings, and (6) own or learned rules. In addition, two error categories were included: a grammatical error category and a handwriting error category.

Results

Spelling Strategies Analysis: The Spelling Process

The presentation of findings is descriptive and focuses on comparisons within and between grades of the spelling strategies employed and between strategies and error types.

An analysis of the graph (Figure 1) illustrating the relative reliance on the various strategies identified by Read (1975) and Harste et al. (1983) indicates that whereas students at all levels rely on a variety of strategies, students in first grade and first-year university each show greater emphasis than other levels on the use of a particular strategy. First-grade students exhibited a strong reliance on the articulatory features of words in their spelling attempts. First-year university students, on the other hand, were found to rely more heavily on the use of spelling rules than did the younger students.

It is instructive to note the error profiles within each grade level. Errors based on point of articulation account for over half (56.19%) of the nonstandard spellings used by the first-grade subjects. The next largest categories of nonstandard spellings for first-grade subjects are spelling based on the way a word looks (14.95%) and on the way the youngster has solved similar spelling problems, for example, the combination of smaller words to spell a larger word as in "sell fish" for "selfish." Clearly, the major influence on these subjects' spellings is point of articulation, which would indicate that use of articulatory features is a strategy which is relied on heavily by children in the early primary grades.

The error profile for subjects in fourth grade is similar in overall shape to that of first-grade subjects, with point of articulation appearing as the most commonly employed strategy (28.53%) followed by strategies based on the way the youngster has solved similar problems (19.21%) and on the way the word looks (17.23%).

The eighth grade marks a turning point in the strategies employed for spelling, as evidenced by a change in the error profile. Subjects at this level employed the strategy of spelling based on the way a word looks in 28.53%

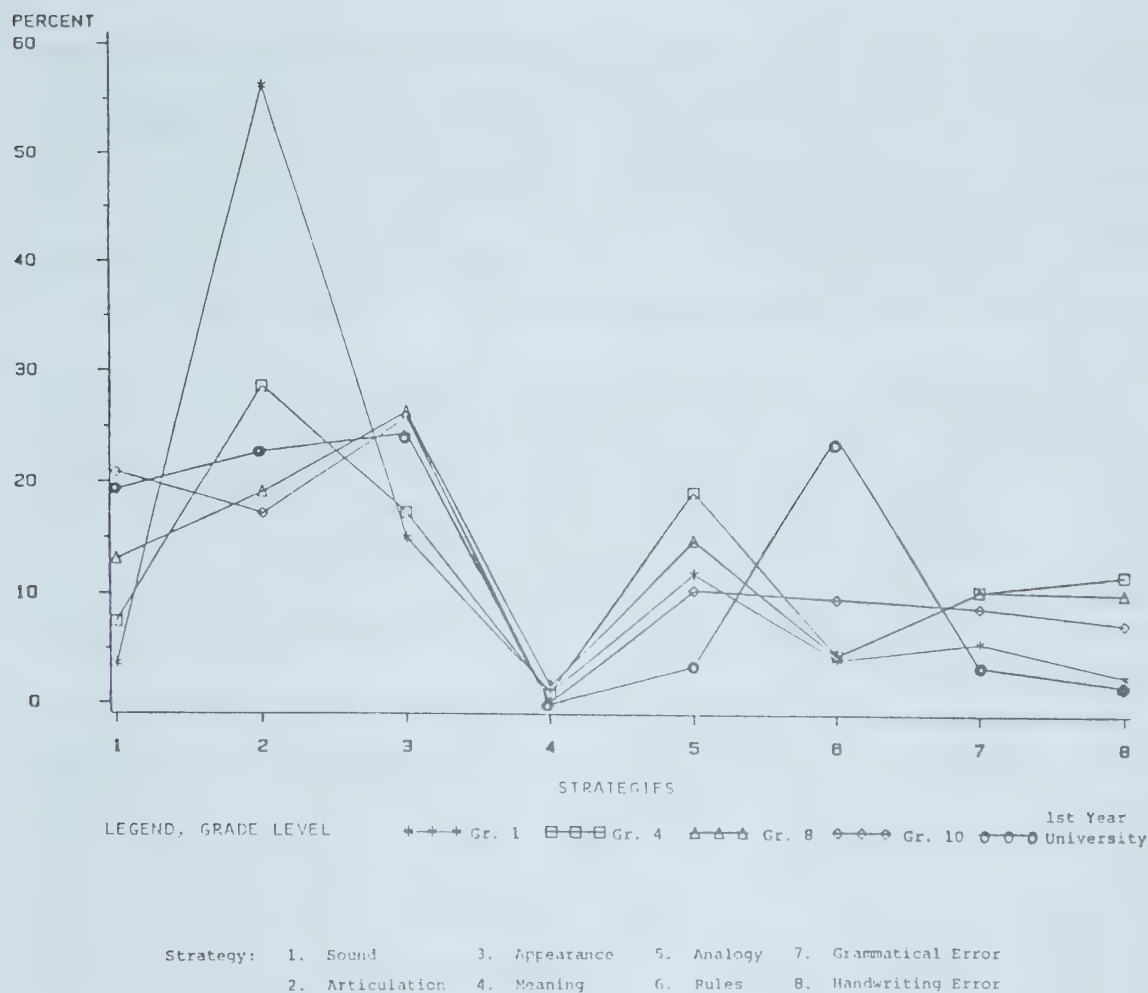


Figure 1: Use of spelling strategies by grade

of the nonstandard spellings. The increased reliance on this strategy (14.95% in first grade, 17.23% in fourth grade, 28.53% in eighth grade) and the concurrent decrease in reliance on articulatory features (56.19% in first grade, 28.53% in fourth grade, 19.03% in eighth grade) appear to indicate a developmental shift toward greater awareness and use of the visual aspects of the orthographic system in solving spelling problems.

The error profile of subjects in tenth grade differs from that of eighth grade, with the greatest reliance still being on spelling based on the way a word looks (25.93%), followed by the way the word sounds (20.74%), and then by articulatory features (17.04%). Students at this level appear to be maintaining their use of both visual strategies and articulatory/phonological strategies. Use of rule-based strategies shows a slight but notable increase from just over 4% in grades 1, 4, and 8 to nearly 10% in tenth grade.

This slight shift can be seen as the beginning of a second major change, as the profile for first-year university students shows greatest reliance on strategies based on rules (24.56%) and on the way a word looks (24.56%), closely followed by use of articulatory features (22.81%) and the way the word sounds (19.30%).

In comparing the grade-level profiles, certain trends become apparent. Between first grade and first-year university, use of strategies 1 (the way it

sounds), 3 (the way it looks), and 6 (my rules or someone else's) increases. Strategy 2 (the way it articulates) drops off sharply between first and fourth grades, and, to a much lesser degree, between fourth and eighth grades. Overall, the most interesting points of change are the shifts away from reliance on articulatory features by the end of the primary years, toward greater use of visual features around the eighth grade, and the increasing incorporation of rule-based strategies between tenth grade and first-year university.

The lists of words which were spelled in nonstandard ways were analyzed in an effort to identify spelling patterns which were particularly problematic for students at various grade levels (Table 1). Computer programs based on Read (1975) were developed to identify and list nonstandard spellings of the following orthographic patterns: blends/clusters (...rm, ...spl...); sonorant consonants (...tle, ...son); diphthongs (...oy, ...oi...); vowel-controlled *r/er* (...ar, ...er); *s/z* sibilants (...us..., ...is...); digraphs (sh..., ch...); preconsonantal nasals (...nt, ...ng).

The most apparent information gained from this analysis is that there are many nonstandard spellings which are not accounted for by these error categories. The 10 categories contain between 21% and 30% of the nonstandard spellings for the subjects in the study. There are, however, a few interesting results based on the error patterns and totals found in each grade level. Diphthongs and vowel controlled *r/er* accounted for over half of the items listed in these 10 categories for first-grade students. Although these two types continued to provide some difficulty through eighth grade, other categories were more prominent after first grade. Fourth grade students exhibited greater difficulty with digraphs than with any other category. Also, this level had a considerably greater number of digraph errors than did any other level.

Subjects in the eighth grade had a surprisingly high number of preconsonantal nasal errors. Despite the overall drop in assessed reliance on articulatory features and Read's description of preconsonantal nasal errors as evidence of such a strategy, the greatest number of eighth-grade subjects' errors (as represented by these 10 orthographic categories) occurred in the preconsonantal nasal category. Despite the marked reliance on articulatory features by the first-grade subjects, preconsonantal nasal errors accounted for only 12% of the errors in the 10 categories, while 20% of the eighth-grade errors in this analysis were in the preconsonantal nasal category. Tenth-grade students exhibited greatest difficulty with blends, as did the first-year university subjects. The latter also had difficulties with the vowel controlled *r/er*.

Error Category Analysis: The Spelling Product

In her 1985 study of spelling errors made by learners of English as a first and as a second language, Bebout devised a comprehensive error categorization system. Her categorization scheme employs eight categories. This system was used in the present study because of its comprehensiveness and its focus on English orthography and also as a contrast to the strategies ap-

Table 1
Errors by Orthographic Patterns

Grade Level	Orthographic Patterns	Sonorant		Vowel		Preconsonantal		GH	TR/DR	NG	# of errors in the 10 categories	% of total grade level accounted for by these 10 categories
		Blends	Consonants	Diphthongs	Controlled R	S/Z	Digraphs	Nasals				
1	No. of errors	2.0	1.0	11.0	12.0	1.0	5.0	5.0	4.0	0	0	21%
	% of errors in these categories	5.0	2.0	27.0	29.0	2.0	12.0	12.0	10.0	--	--	
	% of errors in this grade	1.0	0.5	6.0	6.0	0.5	2.0	2.5	2.0	--	--	
4	No. of errors	14.0	5.0	12.0	14.0	1.0	19.0	15.0	6.0	2.0	2.0	28%
	% of errors in these categories	14.0	5.0	12.0	14.0	1.0	10.0	15.0	6.0	2.0	2.0	
	% of errors in this grade	4.0	1.0	3.0	4.0	0.2	5.0	4.0	1.6	0.5	0.5	
8	No. of errors	12.0	8.0	9.0	14.0	1.0	9.0	17.0	6.0	1.0	6.0	19%
	% of errors in these categories	14.0	10.0	11.0	17.0	1.0	11.0	20.0	7.0	1.0	7.0	
	% of errors in this grade	3.0	2.0	2.5	3.0	0.2	2.5	4.0	1.0	0.2	1.0	
10	No. of errors	11.0	3.0	4.0	6.0	4.0	2.0	3.0	7.0	0.0	1.0	29%
	% of errors in these categories	27.0	7.0	10.0	15.0	10.0	5.0	7.0	17.0	--	2.0	
	% of errors in this grade	8.0	2.0	7.0	4.0	7.0	1.0	2.0	5.0	--	0.7	
UNIV	No. of errors	8.0	3.0	--	7.0	--	--	1.0	1.0	--	20.0	30%
	% of errors in these categories	40.0	15.0		35.0			5.0	5.0			
	% of errors in this grade	12.0	4.5		11.0			1.5	1.5			

proach being employed. This procedure permitted the researcher to compare results from two classification schemes, that is, a process orientation (the Harste et al. and Read categories) and the product orientation (the Bebout categories).

Table 2 presents the results of the overall categorization of errors where the spelling errors were placed in one or more of the general error categories. The information on multiple errors consists of those words containing two different categories of errors.

Distinct and dramatic changes in the kinds of spelling errors made by student writers occurred over the 13-year sample. First-grade students seemed to find vowels the most troublesome orthographic feature (28.9%). To this might be added the *schwa* (11.8%) and the "silent e" (15.3%) which are separately coded vowel error categories. As well, their spelling was categorized as unanalyzable more often than any other group (18.9%). This finding can be coupled with the incidence of multiple errors (21.5%), again the highest for the total sample.

Fourth-grade students had much more frequent difficulty representing consonant sounds than the first-grade students (34.2% and 14.9%), while their vowel problems were somewhat less (21.4% to 28.9%), and their unanalyzable spelling and multiple errors decreased (12.4% from 18.9%, and 14.3% from 21.5%). Problems with the *schwa* decreased sharply (3.2% from 11.8%).

Eighth-grade students showed an error distribution very similar to those of the fourth grade with a further decrease in unanalyzable spellings (from 12.4% to 8.8%) and in multiple errors (from 14.3% and 12.1%).

Subjects in the tenth grade showed fewer errors in consonants than the eighth grade (25.3% to 32.7%) but an increase in unanalyzable spelling (from 8.8% to 10.6%). Otherwise their spelling products were similar to those of the fourth and eighth grades.

The first-year university subjects exhibited remarkably different spelling products from the other subjects. They had much more trouble with consonant doubling than the next highest groups—the eighth and tenth grades (26.8% to 10.6%), much less trouble with other consonants than the tenth-grade group (17.9% to 25.3%) and more "other vowel" trouble than even the first-grade subjects (36.6% to 28.9%). On the other hand, their unanalyzed spelling was the lowest for all subjects (6.3%).

The results of the present study contrast sharply with results obtained by Bebout in regard to the occurrence of consonant and vowel errors. Subjects in this study had relatively more trouble with consonants than did Bebout's subjects except in the first grade and at university level.

Comparison of Process and Product Analyses

For the first-grade students it appeared that point of articulation of a sound is the most important spelling strategy employed. At the same time, an analysis of the orthographic features underlying their spelling errors showed that they have the greatest difficulty with both vowels and "vowel-like" representations (*schwa* and silent *e*). The orthographic representation

Table 2
Error Analysis by Category and Grade (%)

Grade Category Level	Consonant Doubling 1	Other Consonant 2	Schwa 3	Silent <e> 4	Other Vowel 5	Misordered Letters 6	Unanalyzable 7	Multiple Homophones 8	Percentage of Errors
1	6.1	14.9	11.8	15.3	28.9	2.2	18.9	1.8	21.5
4	6.8	34.2	3.2	9.7	21.4	3.6	12.4	8.7	14.3
8	10.6	32.7	1.4	9.0	22.5	4.9	8.8	10.0	12.1
10	10.6	25.3	1.2	9.8	24.9	4.5	10.6	13.1	13.1
UNIV	26.8	17.9	0.9	2.7	36.6	0.9	6.3	8.0	14.3

of the “open” articulation of the vowel sounds appeared to be most problematic for beginning spellers. For English speakers the differences in articulation between different vowels are less extreme than the differences in articulation between consonants. Thus the letters used to represent vowels stand for less distinctive differences than is the case for consonants. These physiological, phonetic, and orthographic factors may engender the observed spelling behaviors of primary students.

Fourth-grade students, while sharing point of articulation as an important strategy with the first-grade subjects, now looked to analogy and appearance as valuable spelling strategies. Their spelling products still showed difficulty with vowels and vowel-like sounds, but consonants emerged at this time as a major spelling problem. Presumably, analogy and appearance provided insufficient specific information about consonant representation in spelling at this level, due, perhaps, to lack of literary experience.

Eighth-grade subjects (like the fourth-grade students) had begun to move strongly toward the use of visual strategies, for example, appearance and analogy, to support their use of articulation and sound values in solving spelling problems. This behavior resulted in spelling products similar to those of the fourth-grade subjects, with consonants emerging as a major spelling problem. The reasons for this outcome appeared to be similar to those for the younger group and showed that eighth-grade subjects had yet to master certain rules and principles of spelling.

Tenth-grade students, while generally maintaining their use of sound, articulatory, and appearance strategies, but lessening their use of analogy, appeared to be finally moving toward the use of rules and principles in spelling. This move was also seen in the general increase in the use of orthographic patterns such as digraphs and those representing diphthongs as shown in Table 2. The increased use of rules appeared to be one of the clearest signs of maturity in spelling behavior and was related to the spelling products of tenth-grade students in that their error pattern showed a wide distribution across error types. They seemed to have mastered many of the sound symbol relationships and the rules that govern English orthography; thus their errors reflected the general possibilities for error rather than focusing on certain types, thereby suggesting lack of knowledge of other possibilities.

At the university level, subjects continued to use sound, articulatory, and appearance information, but they moved dramatically toward the use of rules and principles in spelling. This shift seemed to generate certain spelling product problems in that this group produced a greater proportion of consonant doubling errors (specifically doubling when no doubling is required in nonaffix spellings) than any other group, as well as vowel problems other than the *schwa* and silent *e*. It appeared that these students were applying doubling rules too frequently, perhaps having been alerted to the presence of this rule. Similarly, they may have been relying on rules for the representation of vowels more than on articulatory information with the consequent possibility of error. They had, however, apparently mastered

the various representations of the *schwa* sounds and the rules for the use of silent *e*.

Research Questions

Answers to the five general questions that guided the design of this study summarize the results of this research:

1. Are the findings of Read, Harste et al., and Bebout replicable?

The clear answer is *yes*. Strategies identified by Read and Harste et al. and error categories developed by Bebout were capable of being reliably used to analyze spelling behavior. Results were generally comparable with those of the originators of the schemes with some variation, due presumably to the nature of the subjects and the source of spelling errors—connected discourse rather than individual words.

2. How do the spelling strategies and spelling products change over time?

Spelling strategies of subjects changed from an emphasis on the sound and articulatory features of words to a broad-based approach using a variety of strategies but with special growth in the use of spelling rules. Products changed from greater difficulty with vowels to difficulties with consonants to a general range of errors sampling each of the possible error types. Subjects seemed to master first one then another aspect of English orthography as they developed.

3. Do these changes provide evidence of developmental trends?

The clear answer here again is *yes*. Subjects appeared to move toward a broad based approach to the use of spelling strategies and errors generally distributed over the whole range of possibilities reflecting the statistical possibility of such errors.

4. Are particular types of words especially difficult for learners at various stages?

Specific difficulties include the problems of first-grade children with silent *e*, diphthongs, and vowel controlled *r*; consonant doubling problems of older subjects; and the reemergence of problems with vowel controlled *r* by university level students.

5. What are the relationships between spelling processes and spelling products?

An emphasis on the use of certain strategies such as sound and articulation appears to lead to spelling errors wherein those strategies are manifested. The broader the range of strategies employed, the more general the distribution of spelling errors.

Conclusions

The results of this study confirm that students proceed systematically to develop systems or strategies for representing the sounds of words. This systematic process is seen as a semiotic one where language users learn to employ symbols (in this case orthographic symbols) in a conventional man-

ner. The learner establishes the essential semiotic relationship early and subsequently refines it. This semiotic relationship is the connection between the arbitrary signs of the English alphabetic system and the sound values associated with spoken words. It is evident from the study that only the simplest sort of relationships are developed in the early stages of spelling acquisition and that these relationships are later broadened to include a range of connections between symbol and sound. These varied and complex connections result in a number of strategies used to correctly spell a word. Students thus learn to replace a sound in language with a particular symbol or combination of symbols. The development of this ability appears to proceed much like the development of any other language ability in that the learner identifies basic principles and subsequently elaborates them.

Spelling strategies appear to grow from an emphasis on the sound values and the appearance of a word to the use of this information in conjunction with rules, principles, and generalizations that govern English orthography. The nature of individuals' spelling products (their error patterns) reflects this change. Students, therefore, appear to need exposure to language in its oral and written form, and to the rules that describe standard orthography. In view of the considerable emphasis on the latter in the elementary school and the lack of use of this strategy until the tenth grade and the university level, long after formal teaching of spelling has ended, it appears that teaching rules is an inefficient pedagogical strategy. It may, however, be that individuals take a period of time to internalize these rules and generate their own system of spelling strategies. Therefore, it may be most appropriate to provide learners with much exposure to correct spelling through reading and functional spelling activities and to the sounds of language through speech while encouraging them to discover conventional orthography and the rules that govern it through writing.

Spelling ability appears to evolve in a predictable manner and, while some individuals will always have difficulty with spelling, most learn to spell very well. By the time they have reached university their rate of misspelled words has dropped from an average of about 13% in grade one to an average of about one half of 1% at university.

Suggestions for Further Research

In view of the current emphasis on "Whole Language," attention should be given to determining the effect of this kind of language experience on spelling strategies. As well, the effect of diagnostic approaches to spelling wherein a student's spelling strategies form the basis of instruction should be examined.

References

- Anderson, K.F. (1985). The development of spelling ability and linguistic strategies. *The Reading Teacher*, 39(2), 140-147.
- Bebout, L. (1985). An error analysis of misspellings made by learners of English as a first and as a second language. *Journal of Psycholinguistic Research*, 14(6), 569-593.
- Beers, J., & Henderson, E. (1977). A study of developing orthographic concepts among first graders. *Research in the Teaching of English*, 11, 133-148.

- Campbell, D.T., & Stanley, J.C. (1963). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally.
- Ehri, L. (1980). The development of orthographic images. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 311-338). New York: Academic Press.
- Frith, U. (1980). Unexpected spelling problems. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 495-515). London: Academic Press.
- Gentry, J.R. (1978). Early spelling strategies. *Elementary School Journal*, 34, 88-92.
- Gentry, J.R. (1984). Developmental aspects of learning to spell. *Academic Therapy*, 20(1), 11-19.
- Harste, J.C., Burke, C.L., & Woodward, V.A. (1983). *The young child as writer-reader, and informant* (Report No. NIE-G-80-0121). Bloomington, IN: National Institute of Education.
- Read, C. (1975). *Children's categorization of speech sounds in English* (NCTE Research Report No. 17). Urbana, IL: National Council of Teachers of English.
- Zutell, J. (1980). Children's spelling strategies and their cognitive development. In E. Henderson & J. Beers (Eds.), *Developmental and cognitive aspects of learning to spell: A reflection of word knowledge* (pp. 52-73). Newark, DE: International Reading Association.

JOE BELANGER

University of British Columbia

R. GLENN MARTIN

University of Alberta

The Influence of a Reading Treatment on the Written Compositions of Grade 11 Students: A Research Update

The research reported below replicates on the grade 11 level our earlier study of the effect of a reading treatment on the reading achievement and writing proficiency of a sample of students in grades 9 and 10 (Belanger & Martin, 1984) and alters the factor we deemed responsible for the lack of writing improvement in the earlier study. In the current study writing was taught systematically and evaluated as part of students' regular classroom work in English. Sixty-three experimental and 48 control students wrote both the pretest and posttest (Nelson-Denny Reading Test) and in-class essay test. Analysis of covariance indicates significant differences favoring the experimental group on both the reading test ($p=0.046$) and the writing sample ($p=0.003$) four months following the administration of the treatment. Final tests will be administered at the end of the following year, 16 months after the experiment began.

Background

Although correlations between reading achievement and writing proficiency are well known and well documented, a 50-year search for a reading treatment which improves writing skill has until now been unsuccessful. Moderate, statistically significant correlations have been reported between general reading and writing ability (Loban, 1976), reading and the syntactic complexity of writing (Hartman, 1984), reading ability and correct spelling and editorial usage (Pitts, 1985), reading and environmental factors (Weathermon, 1984), and the teaching of beginning reading and writing ability (Eckhoff, 1985). In a meta-analysis of 89 studies conducted between 1950 and 1978 which correlated reading, writing, speaking, and listening measures, Hammill and McNutt (1980) reported that the median correlation between 37 measures of reading and writing was .61. However, because empirical studies have not been able to show a causative relationship between reading and writing skills, these correlations have been poorly understood.

Previous reading treatments have not improved writing ability. Treatments which have taught vocabulary development and paragraph reading

Joe Belanger is an assistant professor in the Language Education Department at the University of British Columbia.

R. Glenn Martin is professor emeritus of education at the University of Alberta.

skills (Eurich, 1931), the SRA materials (Schneider, 1970), psycholinguistic methods of teaching remedial reading (O'Donnell, 1974), or stylistic and rhetorical principles (Austin, 1983) failed on the first step—the improvement of reading—and therefore conclusions could not be drawn about the hypothesized consequent growth in writing ability. An earlier study by the authors of this paper (Belanger & Martin, 1984) reported that a phonics-based reading treatment resulted in significant ($p < 0.0013$) and substantial (about one half year's growth on the test norms) differences on a standardized reading measure but no differences on a writing sample. We conjectured that a lack of incentive for the students to write and a lack of adequate instruction in writing prevented the subjects from producing their best written work; experimental and control essays were uniformly poor.

A small number of writing treatments which have taught aspects of the structure of English appear to have the potential of improving both reading and writing skill, but only studies using sentence combining on the elementary level have produced completely unambiguous results. Sentence combining studies by McAfee (1980, grade 5) and Mackie (1982, grade 4) resulted in significant improvements on both reading and writing measures. However, a number of other studies of sentence combining at the elementary level have produced significant differences on some reading measures but not on others (Combs, 1977; Levine, 1976; Straw & Schreiner, 1982). On the grade 6 level, Kelley (1984) reported that teaching the structure of paragraphs and stories produced significant results on both reading and writing measures when compared with free reading, but other studies which have taught aspects of the structure of English have produced ambiguous results. Taylor and Beach (1984), Crowhurst (forthcoming), Braun and Gordon (1984), and Bossone and Troyka (1976) reported significant changes in writing ability but inconsistent changes in reading ability as a result of the study of various schemata and text-structure treatments. (See Belanger, 1987, for a summary of these and other studies.)

Design and Procedures

A true random pretest/posttest design with a midtest four months following the administration of the treatment was used in this experiment. True randomization was possible because teachers allowed groups of students to be withdrawn from each class while control students worked on individual assignments. The design can be depicted as follows:

R O₁ X O₃ O₅

R O₂ .. O₄ O₆

Findings from the first four months (O₃ and O₄) are reported below.

The 164 grade 11 students who wrote both the Nelson-Denny Reading Test (Form C) administered by the English Department in the fall and the department-wide Fall Essay Test were randomly designated experimental or control. In February, the treatment groups were administered the reading treatment, a four-page device designed to tap students' phonic knowledge (knowledge of the code of English) to enable them to rely on it to decode the printed page.¹ Five students, three experimental and two control,

who learned English as a second language after enrolling in grade one, and three experimental students who missed the treatment due to long-term illness were eliminated from the experiment. In June, after eliminating school dropouts and transfers (25) and those missing either the essay (9) or the reading (19) tests, 111 students (63 experimental and 48 control) remained.

The Nelson-Denny Reading Test, Forms C and D, were used as the reading measure in the experiment. Form C was administered by classroom teachers in September 1986 as part of regular English Department data collection procedures. Form D was administered in June by classroom teachers for the experiment. Twelve months following the administration of Form D, Form A will be used as a posttest.

The writing sample consisted of two essays written during normal school testing periods as part of an assessment in English and marked by the classroom teachers during released time using procedures suggested by Diederich (1974). (School officials agreed to release teachers to score the essays if substitutes were hired to teach the regular classes.) Papers were identified by a code number only, and examinations from the five classes were randomly intermingled prior to the marking sessions. Raters, therefore, were unaware of the identities of the authors. One of the investigators (Belanger) joined the marking team. Teachers were accustomed to team marking because the department had, on occasion, taken a day to assess student writing as a team. In addition, four of the six evaluators had served together on provincial marking teams. Following Diederich's suggestions, essays were assigned percentage scores by two primary raters. If these scores were nine marks or fewer apart, the paper's grade was an average of the scores of the two primary raters. Reliabilities for the two primary raters (unadjusted scores) were .53 in the fall and .70 in the spring. If the scores of the primary raters differed by 10 marks or more, a third rater was assigned the paper to rerate without knowledge of the scores of the other two raters. Elimination of the discrepant grade resulted in interrater reliabilities of .85 in the fall and .88 in the spring. Twelve months following the collection of the second writing sample, a third essay (written and evaluated under conditions similar to those used for the first two) will be used as a posttest.

All statistical procedures were carried out on the UBC Amdahl V8 computer using the Statistical Package for Social Sciences Extended Version (SPSSX, Release 2.0). Two-way analysis of covariance (ANOVA) was selected to test major differences between the experimental and control groups because the pretest revealed differences (although nonsignificant) between the two groups on both the initial reading and writing tests (Table 1).

Interactions between the treatment and both the sex and the reading and writing achievement of the subjects were tested. To establish reading and writing achievement levels, students were divided into four categories based on their combined initial reading and writing scores. As Table 2 shows, cell sizes for ability level ranged from 11 to 17 subjects.

Table 1
Raw Scores by Test Period and Treatment

		Fall		Spring		Difference
	n	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$
<i>Reading</i>						
Experimental	63	62.67	22.1	70.73	22.4	8.06
Control	48	58.90	18.6	62.44	21.9	3.54
Exp. Girls	34	67.41	24.1	75.03	23.0	7.62
Exp. Boys	29	57.10	18.4	65.69	20.9	8.59
Con. Girls	24	62.21	21.5	65.88	25.6	3.67
Con. Boys	24	55.58	15.0	59.00	17.2	3.42
<i>Essay</i>						
Experimental	63	61.38	9.9	65.57	11.3	4.19
Control	48	63.10	9.2	61.60	8.4	-1.50
Exp. Girls	34	64.60	8.7	68.85	10.0	4.25
Exp. Boys	29	57.59	9.9	61.72	11.8	4.13
Con. Girls	24	63.83	8.2	64.08	9.4	0.25
Con. Boys	24	62.38	10.3	59.13	6.4	-3.25

Interrater reliabilities and correlations between reading and writing ability were calculated using PEARSON CORR (zero-order product moment correlations).

Results

Reading

Analysis of the reading measure corroborates the findings of our earlier study (Belanger & Martin, 1984). Four months following the administration of the treatment, using the pretest as the covariate, the experimental groups were superior to the control groups at the .046 level of confidence on the June reading test (Table 3).

As Table 1 indicates, the experimental and control students experienced substantial growth on the Nelson-Denny Reading Test over the four-month initial phase, but the experimental students (who gained 8.06 raw-score points or 12.9%) showed over double the gains of the control students (who gained 3.54 points or 6.0%). This difference represents about one half year on the test norms. This pattern is similar to that of our 1984 study which was conducted at the grade 9 and grade 10 levels and used the Davis Reading Test.

Tests of interactions (Table 3) indicate that neither the sex of the subject nor the subject's initial reading and writing ability influenced the results

Table 2
Raw Scores by Test Period, Treatment, and Ability

		Fall		Spring		Difference
	n	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$
<i>Reading</i>						
Low Exp.	17	42.00	7.2	53.00	14.4	11.00
Con.	11	41.46	9.2	49.55	15.6	8.09
2 Quart. Exp.	15	52.60	6.5	60.07	12.0	7.47
Con.	13	51.85	5.8	56.54	9.8	4.69
3 Quart. Exp.	14	62.21	8.5	70.57	11.4	8.36
Con.	13	60.69	6.9	57.31	9.3	-3.38
High Exp.	17	92.58	15.7	98.00	16.6	5.42
Con.	11	82.55	20.8	88.36	28.0	5.81
<i>Essay</i>						
Low Exp.	17	53.94	8.9	61.24	9.3	7.30
Con.	11	53.00	5.7	55.09	6.4	2.09
2 Quart. Exp.	15	59.07	6.9	59.33	8.2	0.26
Con.	13	59.85	5.4	60.54	6.7	0.69
3 Quart. Exp.	14	64.14	5.9	69.21	12.6	5.07
Con.	13	68.54	7.9	63.23	5.7	-5.31
High Exp.	17	68.59	10.1	72.41	10.3	3.82
Con.	11	70.64	5.5	67.46	10.3	-3.18

significantly. As Table 1 shows, the experimental boys showed somewhat more growth (8.59 points) than the experimental girls (7.62 points), while the control girls made slightly greater gains than the control boys. There were also small differences when the groups were divided on the basis of initial reading and writing ability. Table 2 reveals that all but the top quartile of experimental students showed reading gains greater than those of the control students. In our 1984 experiment, on the other hand, nonsignificant differences favored the girls and the high-ability students.

Results

Writing

Analysis of the students' essays revealed significant differences between the experimental and control groups, unlike the analysis in our 1984 project. As Table 4 shows, when the experimental and control groups are compared on the June writing sample using the pretest as covariate, the experimental students are superior to the control at the .003 level of confidence. Table 1 reveals that the experimental students gained 5.69 points over the control students. The experimental students gained 4.19 points while the control groups lost an average of 1.5 points.²

Table 3
Spring Reading Score Using Fall Reading Score as Covariate

Source	DF	Mean Square	F	Prob
Covariate	1	41181.28	325.21	0.0001
Main Effects	1	527.87	4.08	0.046
Interactions				
Sex	1	12.02	0.09	0.761
Ability	3	180.05	1.39	0.250
Explained	13	3301.87	25.51	0.0
Residual	97	129.42		
Total	110	504.34		

As with the reading results, there were no significant interactions between the treatment and the sex or the ability of the subjects (see Table 4). As Table 1 shows, the experimental boys and girls showed almost equal pre/post gains (4.13 and 4.25 points respectively). The control girls, on the other hand, showed only slight gains while the control boys accounted for all the losses shown by the control groups. When the groups are examined by ability (Table 2), all but the second quartile experimental students show substantial writing gains over the control students. Considering the students' writing, then, the treatment did not appear to offer a particular advantage to boys or girls or to weak or strong students.

Discussion and Implications

In 1984 we reported that significant ($p < .0013$) and substantial (almost one half year's growth on reading-test norms) gains by experimental students on a standardized reading test were not accompanied by measurable

Table 4
Spring Essay Score Using Fall Essay Score as Covariate

Source	DF	Mean Square	F	Prob
Covariate	1	3912.96	59.73	0.0001
Main Effects	1	576.56	9.14	0.003
Interactions				
Sex	1	26.29	0.42	0.520
Ability	3	119.76	1.90	0.135
Explained	13	427.98	6.78	0.0001
Residual	97	63.11		
Total	110	106.23		

changes on writing measures. The changes on the writing measures were not, in fact, generally in the hypothesized direction. Measures taken both three and six months following the administration of the treatment showed that the major effect of the reading treatment was evident in the first three months of the experiment and that the gap between the experimental and control students continued to widen, albeit more slowly, during the second three months. The writing results, on the other hand, were random and erratic at both test periods.

We felt in our 1984 study that a major difficulty with our findings was the lack of incentive to write provided by the data collection procedures. Students were asked to write an essay in October, February, and June during their regular class periods but without preparation, response, or evaluation. The essays were neither marked by their teachers nor returned. Although this type of demand assignment with its lack of preparation, audience, response, and publication has been one of the hallmarks of research on writing, there is little reason to expect that students will produce their best writing—writing which would accurately reflect their ability—under such conditions. Nor did writing appear to be a major part of the curriculum in the classrooms studied. As is the case in many English classes (Applebee, 1981), writing was sometimes expected but seldom taught.

Consequently, for the current experiment we sought a setting in which writing was taught systematically and valued as an important form of communication. In addition, the writing samples were required to be written as part of the students' regularly assigned classroom work, marked by teams of teachers, and counted as a portion of the students' marks in English class. Teachers in the department agreed to assign essays as term tests in writing, administer them under examination conditions, and score them on a day released from school. There were also three minor differences between the current experiment and the 1984 experiment, differences which were not expected to influence the findings: the grade level (9 and 10 in 1984; 11 in the current experiment); the size of the experimental population (198 then; 111 now); and the standardized reading test (Davis in 1984; Nelson-Denny currently³).

We believe these results are important because they answer a question that has intrigued researchers for at least half a century: Does an improvement in reading ability bring about improvement in writing ability? Previous research has been able to demonstrate only a correlational relationship between reading treatments and writing achievement, a relationship which could be attributed to a variety of causes. The current study, on the other hand, shows that a reading treatment has a direct influence on writing ability by establishing a causal link between the two, a link which suggests an explanation for the vast amount of correlational data in earlier research. As such, the findings shed new light on the theoretical bases of reading and writing.

The major question raised by the findings is, What characteristics of the treatment are responsible for the observed changes on the reading and writing measures? As our outline of the conceptual basis of the reading treat-

ment (Note 1) suggests, we attribute the results of the study to the students' improved automatic handling of the code of written English.

Obviously this is speculative inference from the data, not direct observation. But because the results offer little comfort to prevailing theory in the field of reading, some explanation is in order as to why we undertook this study and its predecessors and what theoretical implications we see as accounting for results that seem to challenge conventional wisdom. What follows is our interpretation of the results, along with the direction in which they seem to point.

On the reading measure, gains resulted, we believe, from the students' improved ability to decode written words automatically and fluently. The treatment forced students to rely on their ability to decode the sequence of sounds embodied in the written word, an ability students have apparently acquired through nine or more years of exposure to print in the school system and elsewhere.

Much current reading instruction (such as word guessing from context), however well intentioned, may have the effect of interfering with induced capacity by introducing a conscious element in a process best carried out automatically. It is this interference which we believe is counteracted by the SOS (Sequence of Sounds) technique, and if the interference from counterproductive instruction is removed, preexisting but unused capacity can manifest itself.

The effect on writing, on the other hand, while also a direct result of improved ability to handle the written code of English, is probably more subtle and based on a variety of benefits produced by improved (or newly liberated) facility in automatic decoding during reading. Changes on the writing measures are likely the cumulative effects of a number of separate aspects of handling the written code, each of which might be enhanced by formal instruction, with improved reading ability providing the foundation on which such instruction might prosper.

On the word level, improved handling of the written code may lead to increased facility with English spelling and, in due course, to broader vocabulary. Students taught to guess or memorize words when they are reading are likely to rely on guesswork and memory in their writing. On the other hand, students who have cracked the code and who become proficient in using the English spelling system can rely on the code for the many regular English words, thus reducing the memory load to those words, or word portions, which do not fit the system. This, of course, is not a change that comes overnight, but one which takes place gradually as writers figure out the system; nonetheless, small improvements might be expected in the beginning, growth which, when taken cumulatively with other developments, results in significant overall improvement in writing ability.

Vocabulary is, we believe, amenable to similar kinds of enrichment. Students who have cracked the code and learned to "hear" words accurately as they read them—rather than to guess what the word might be—know when they have encountered an unfamiliar word. Having perhaps heard the word once or twice, they can decide whether they want to add it to their

vocabularies. However, readers who rely on guessing rather than the code have two variables to try to control: first, whether they have guessed correctly what the word *is* (that is, if they have guessed correctly what spoken word has been printed), and second, whether (now having guessed what the word is) they can proceed to guess what it means. For example, in the sentence “the boy was adamant about his role in the theft of the candy,” the decoder hears the sequence of sounds and knows whether “adamant” is part of his vocabulary. The word guesser, on the other hand, does not know whether the word is *adamant*, *abhorrent*, *aberrant*, or what, and therefore reads “the boy was *something* about his role....” Unless he returns to *adamant* to decode it (or asks his teacher what the word is) he cannot know whether the word *adamant* is in his vocabulary because he has not read *adamant* but merely glossed over it or guessed what kind of word might fit in *adamant*’s position. If he has not read the word, he does not know if he knows it; if he does not know whether he knows it, he has no chance of either understanding or learning it.

On the sentence level, automatic decoding takes the emphasis off the word and puts it on the phrase, clause, or sentence. This gives the student the opportunity to observe and absorb the cadence and rhythm of written English, as well as the sense supplied by syntax that is different from the sum of the particular word meanings. The word guesser, on the other hand, skipping some words completely and returning to others to puzzle them out, cannot hear the ebb and flow of the language, and the strength of the syntax of English will not be absorbed. Reformed word guessers who have been taught to rely on the code, on the other hand, can benefit from listening to the rhythms of the page, and can slowly begin to develop a sense of the metrics and logic of written English. But especially beyond the sentence, there will be little flow of meaning for the unreformed word guesser who must read from guess to guess, with little time for the larger issues in the text.

One of the enduring theoretical bases for expecting relationships between reading and writing has been the observation that reading provides ideas and models for writing. Under this assumption, one who reads widely gains a command of things to write about and a notion of how writers organize, provide evidence, and nurture readers. Even if improved reading did not automatically lead to more reading, it would increase what students learn from the reading they do. As with spelling, vocabulary, and syntax changes discussed above, improvement in ideas and style probably contributed some measure to the significant improvement in writing ability noted in the study.

Similar processes appear to be operating on the level of standard editorial usage. Readers who decode automatically and hear the rhythm and flow of the language are likely to hear the need for punctuation. On the other hand, readers who concentrate on guessing what words might fit into a given sentence slot are likely too preoccupied with error reduction to have much time to sort out punctuation.

We believe that all of the above contribute in some way to the improved writing which was brought about by the foundation laid by the reading treatment. While some of these changes in writing might be expected to be automatic, it seems more likely, especially in light of our 1984 results, that they can be enhanced by instruction. Our findings suggest that automatic decoding is basic to such instruction.

The reading results simply confirm findings of a number of earlier experiments using the reading treatment (Belanger & Martin, 1984; Martin, 1975a, 1975b). Taken by themselves, these results have implications for both schools and research. First, we should question the current complacency about the reading competence of average and above average secondary students. Both categories of students show rapid and substantial improvement in reading in this study, suggesting that learning to rely on their knowledge of the code of written English is not the exclusive domain of those traditionally stigmatized as *remedial*. The implication is, of course, that even the most able readers are in some way impaired readers. We believe, therefore, that educators should consider applying the technique across the board to secondary students, not just to those who obviously pose a problem.⁴

Second, the reading results raise questions about the fundamental basis of currently accepted reading theory. By its very nature, the reading treatment runs athwart much of this accepted theory including the orthodox tradition from Gray, Gates, Harris, Robinson, and others and the more recent uncertainty-reduction views of Goodman, Smith, and like-minded theorists. Similarly challenged are secondary reading assumptions leading to techniques that add to the reader's self-consciousness about reading, which in turn may diminish the capacity for automatic decoding.

The study also raises for future research these questions:

1. What kind of instruction in writing is most productive following reading improvement (e.g., sentence structure, voice, coherence, ideas, vocabulary, editorial usage, or spelling)? Is the grade level of the subjects a factor in the effectiveness of the instruction offered?
2. Is the mode of writing a factor? This experiment focused on the expository essay because the grade 11 teachers in the study traditionally taught the expository essay as a major part of their curriculum. The expository essay may or may not be the most suitable mode to teach to capitalize on reading improvement.
3. What are the long-term effects of this reading treatment on writing ability—after five or six years, for example?
4. If a reading treatment that explicitly ignores and bypasses instruction in *meaning* can improve comprehension and enrich the semantic powers of students in their writing, what implications does this have for meaning-based reading theory?

Acknowledgments

The first year of this research project was supported by Social Sciences and Humanities Research Council Leave Fellowship 451-86-0026 and the second year by University of British Columbia HSS Grants Committee Award S87-030.

Notes

1. The references cited below are essential to even a minimal understanding of the SOS procedure and rationale. But the essence of the rationale is suggested in this passage by the inventor of the technique:

All the SOS does is to remove from the conscious mind any consciousness in the activity of reading, except the consciousness of the messages presented in print, exactly as in hearing. Then the mind can think about the thoughts presented while reading. Thus the technique, far from teaching people how to read, stops them from not reading, and allows the subconscious to do what it has been able to do since junior high school, or at least the ninth grade, translate code into sound into word into sentence into paragraph and so on, while the reader can think about the information he is receiving. (Martin, 1976, p. 12)

The procedure involves having subjects read aloud first some single nonsense syllables, followed by polysyllabic units, followed by a mixture of simple words and nonsense syllables, all pronounced accurately without being read aloud by the administrator, followed by short, regular English words which are (under suitable incentives) read much too fast to permit total accuracy, so that conscious control is relinquished and subconscious competence can assert itself. This is preceded by a paragraph of ordinary silent reading and followed by the next paragraph in the same material to demonstrate to the reader that a change has occurred. The procedure is administered optimally to a group of four people, once only, in less than an hour, but can be done successfully with an individual or a group smaller than four. Additional and indispensable aspects of procedure, along with their theoretical bases, are described in Martin (1976, especially chap. II, pp. 22-33) and more casually in Martin (1977). Belanger (1978, pp. 250-53) offers a brief description of the treatment.

Beginning reading programs which share a similar theoretical orientation include the Lippincott series and the DISTAR program.

2. The two essay tests were not evaluated simultaneously; the fall essays were evaluated in the fall and returned to the students, and the spring essays were marked in June. We suspect that the raters were more demanding in the spring (after they had taught the essay form) than they were in the fall. This might account for the apparent losses by the control groups on the second set of essays. The reading tests, on the other hand, were parallel forms of equal difficulty; consequently, increasing demands of raters would not be a factor in the grades.
3. The Davis Reading Test was selected for the 1984 project in part because the school district used it as a system-wide test in grade 9. However, we discovered that by the end of grade 10 many students achieved almost perfect scores on the test. Therefore, we decided that the ceiling on the Davis test would be too low for grades 11 and 12 and chose the Nelson-Denny Reading Test which appeared to be more discriminating for upper-ability students.
4. The nature of the technique limits expenditure to training of staff and about one hour of time for each group of four students in a single, once-only administration. (No commercial materials are required.) Teacher/administrator time and minor adjustments to class schedules are needed to make this possible.

A problem not addressed in this study is that not all willing teachers may be able to give the technique because of its exacting demands on the administrator. Two to three days of instruction in its use appear necessary, along with an open mind and the making of tapes of sample administrations for suitable feedback to prospective administrators of the technique.

References

- Applebee, A.N. (1981). *Writing in the secondary school: English and the content areas* (NCTE Research Report No. 21). Urbana, IL: National Council of Teachers of English.
- Austin, D.E. (1983). *Reading to write: The effect of the analysis of essays on written skills in college composition classes* (Doctoral dissertation, University of Alabama). *Dissertation Abstracts International*, 45, 771A.

- Belanger, J.F. (1978). *Reading skill as an influence on writing skill*. Unpublished doctoral dissertation, University of Alberta. (ERIC Document Reproduction Service No. ED 163 409)
- Belanger, J. (1987). Theory and research into reading and writing connections: A critical review. *Reading-Canada-Lecture*, 5(1), 10-21.
- Belanger, J., & Martin, R.G. (1984). The influence of improved reading skill on writing skill. *Alberta Journal of Educational Research*, 30(3), 194-212.
- Bossone, R.M., & Troyka, L.Q. (1976). *A strategy for coping with high school and college remedial English problems*. (ERIC Document Reproduction Service No. ED 130 268)
- Braun, C., & Gordon, C. (1984). Writing instruction as a metatextual aid to story schema applications. In J.A. Niles (Ed.), *Changing perspectives on research in reading/language processing and instruction* (pp. 61-65). 33rd Yearbook of the National Reading Conference. Rochester, NY: National Reading Conference.
- Combs, W.E. (1977). Sentence-combining practice aids reading comprehension, *Journal of Reading*, 21, 18-24.
- Crowhurst, M. *Reading/writing relationships: An intervention study*. Manuscript submitted for publication.
- Diederich, P.B. (1974). *Measuring growth in English*. Champaign, IL: National Council of Teachers of English.
- Eckhoff, B. (1985). How basal reading texts affect children's writing (Doctoral dissertation, Harvard University). *Dissertation Abstracts International*, 46, 2917A.
- Eurich, A.C. (1931). *The reading abilities of college students: An experimental study* (College Program Series). Minneapolis: University of Minnesota Press.
- Hammill, C., & McNutt, G. (1980). Language abilities and reading: A review of the literature of their relationship. *Elementary School Journal*, 80, 269-277.
- Hartman, D.M. (1984). An investigation into the predictive relationship of ten writing assessment variables to reading comprehension (Doctoral dissertation, University of Connecticut). *Dissertation Abstracts International*, 45, 1353A.
- Kelley, K.R. (1984). The effect of writing instruction on reading comprehension and story writing ability (Doctoral dissertation, University of Pittsburgh). *Dissertation Abstracts International*, 45, 1703A.
- Levine, S.S. (1976). The effects of transformational sentence-combining exercises on the reading comprehension and written comprehension of third-grade children (Doctoral dissertation, Hofstra University). *Dissertation Abstracts International* 37, 6431A.
- Loban, W.D. (1976). *Language development: Kindergarten through grade twelve* (NCTE Research Report No. 18). Urbana, IL: National Council of Teachers of English.
- Mackie, B.C. (1982). The effects of a sentence-combining program on the reading comprehension and written composition of fourth-grade students (Doctoral dissertation, Hofstra University). *Dissertation Abstracts International*, 42, 4779A.
- Martin, E.B.H. (1976). *Administration of the SOS Reading Technique*. Edmonton, AB: Author.
- Martin, R.G. (1975a). *Improving reading comprehension*. Report to the Edmonton Public School Board (duplicated).
- Martin, R.G. (1975b). *SOS Reading Technique data summary from McDougall Junior High School Davis Test score differences, 1974-75*. Report to the Edmonton Public School Board (duplicated).
- Martin, R.G. (1977). More basic than basics. *Arizona English Bulletin*, 20(1), 28-33.
- McAfee, D.C. (1980). Effect of sentence-combining instruction on the reading and writing achievement of fifth grade children in a suburban school (Doctoral dissertation, Texas Woman's University). *Dissertation Abstracts International*, 42, 156A.
- O'Donnell, J.F. (1974). An experimental study of the effects of the supplemental use of a psycholinguistic remedial tutorial program on the reading and writing behaviors of black, high-risk college freshmen, and on their attitudes toward reading, writing, and other college-related stimuli (Doctoral dissertation, Temple University). *Dissertation Abstracts International*, 57, 1552A.
- Pitts, S.J.K. (1985). Spelling as a correlate of reading ability in underprepared college freshmen: Measures and error types (Doctoral dissertation, University of New Mexico). *Dissertation Abstracts International*, 45, 2358A.

- Schneider, V.L. (1970). A study of the effectiveness of emphasizing the teaching of reading skills to improve composition skills in remedial English classes at Kansas City Community Junior College (Doctoral dissertation, University of Kansas). *Dissertation Abstracts International*, 31, 6369A.
- Straw, S., & Schreiner, R. (1982). The effect of sentence manipulation on subsequent measures of reading and listening comprehension. *Reading Research Quarterly*, 17, 339-352.
- Taylor, B., & Beach, R. (1984). The effects of text structure instruction on middle-grade students' comprehension and production of expository text. *Reading Research Quarterly*, 19, 132-146.
- Weathermon, C.J. (1984). *The home backgrounds of writers*. Unpublished master's thesis, University of British Columbia.

JUDITH O'SHEA

Lethbridge Community College

and

KATHERINE BEATY CHISTE

University of Lethbridge

The Topic Selections and Performance of Adult Student Writers

In a one-year study, 249 adult writers of the Alberta Universities' Writing Competence Test were found to differ from younger writers in their performance and topic selections. Older writers performed markedly better than younger ones in terms of grade received. While different topics appealed more to adult writers than to younger writers, there was no correlation for either group between the popularity of a topic and the success of those choosing it. The study concluded that writing topics need not be specifically targeted toward adult writers, nor need personal topics be avoided, as these writers wrote with comparative success across a wide range of topics.

Introduction

Adult students comprise a growing proportion of college and university student populations throughout Canada. At the University of Lethbridge, for example, students aged 25 and over made up 28.5% of the full-time student population in the fall of 1987 (University of Lethbridge, 1988) compared with 16.3% in the fall of 1972 (University of Lethbridge, 1973). When the large numbers of adult part-time students are also counted, it becomes clear that adult learners form a "new majority" of university student populations (Campbell, 1984); administrators and teachers in postsecondary institutions can no longer assume that their students will have the academic experience and preparation of recent high school graduates. In fact, evaluation of student performance becomes more problematic when students vary greatly in academic background. A particularly difficult problem in evaluation is a test of writing competence—another increasingly common phenomenon in universities and colleges.

The writing of adults, in terms of performance and composition process, has been studied to only a limited degree and primarily for purposes of contrast with younger writers (Odell, Cooper, & Courts, 1978). Such studies usually define adults as older teenagers or college students. In examining adult attitudes toward writing, Connors (1982) found that adult students

Judith O'Shea has taught composition at the University of Lethbridge and Lethbridge Community College and is currently an M.Ed. student at the University.

Katherine Beaty Chiste also has taught composition at the University of Lethbridge and is a doctoral student at the University of Calgary.

lacked confidence as writers, spent more time revising essays, and asked for more direction and limits on writing assignments than younger students. Hayes and Flower (1986) studied the planning process of older students and adults, a process characterized more by the use of problem solving strategies than by the use of known writing formats. Shook (1985) argued against setting "personal" writing topics for adults, for whom the self is "no longer the only real source of information or even of interest" (p. 205). Such studies raise questions about topic design, particularly on writing exams where a wide age range of students can be expected; yet it appears that no one has examined comparatively the patterns of topic selection and performance of adult and younger writers.

In 1982 Hoetker commented, "Research to date on topic effects is inconclusive" (p. 378), and in key areas of topic research this is still true. On a Teaching Certificate exam with multiple writing topics, the popularity of an individual topic has been found to be unrelated to the mean scores for that topic (Brossell, 1986). In contrast, Gee (1987), working with grade 12 exams, found that popularity varied inversely with score. Carlman (1986), also with grade 12s, discovered no significant differences between mean scores on topics within a single "mode" (e.g., argument, narration), but did note significant performance differences between modes. A study of holistic essay marking (high school level) found that an argumentative topic led to a greater variety of responses than did a descriptive one and hence perhaps to a greater spread in holistic marks received (Nagy, Evans, & Robinson, *in press*). Researchers have found the subject matter of questions not to be a variable in score received by university sophomores (Brossell & Ash, 1984); others suggest that subject matter, as it relates to topic-specific background knowledge, is a factor in performance (for grade 10 students, Langer, 1983). The delicate question of writers' emotional responses to topic and subject matter has been investigated by Brand and House (1987), who concluded that topic does affect the "emotional fluctuations" and performance of (college) student writers.

The purpose of this study is to examine the topic selections and performance of adult writers (aged 25 and over) of the Alberta Universities' Writing Competence Test, administered since 1983 at the four universities in the province. The AUWCT is a compulsory, timed (two-hour) test in which students respond to one of four questions of general interest in an expository or argumentative essay of about 400 words, expository questions being slightly more common. The four possible grades are Satisfactory and Marginally Satisfactory (passing grades), Marginally Unsatisfactory and Unsatisfactory (failing grades); the marking is detailed rather than holistic, that is, of the "error count" kind.

Our questions are the following: How did the older writers fare on the AUWCT, and how did their performance compare with that of younger ones? What topics appealed or did not appeal to adults and to youth? And finally, did topic choice seem to make any difference in performance?

Methodology

This study draws on the AUWCT exams written at the University of Lethbridge during a one-year period (July 1986 to June 1987) comprised of a total of 1,175 exams written on ten dates. From this total we excluded papers identified by the markers as “English as a Second Language” papers (147 exams). The remaining papers were divided into “youth” exams written by students aged 20 and under (706 papers), “young adult” exams written by students aged 21-24 (73 papers), and “adult” exams written by students aged 25 and over (249 papers). Our research sample of 249 adult writers had a mean age of 32.4; an age profile of these writers appears in Figure 1.

Our data were gathered from the exams themselves, which were stockpiled at the university along with the marking sheets. We recorded the marks achieved (S, MS, MU, or U) by adult and youth writers over the year; for purposes of greatest contrast we discounted the 73 young adult papers. We went on to record the question selections of adult writers on the 10 dates, that is, how many adults chose each question on each date. Using chi-

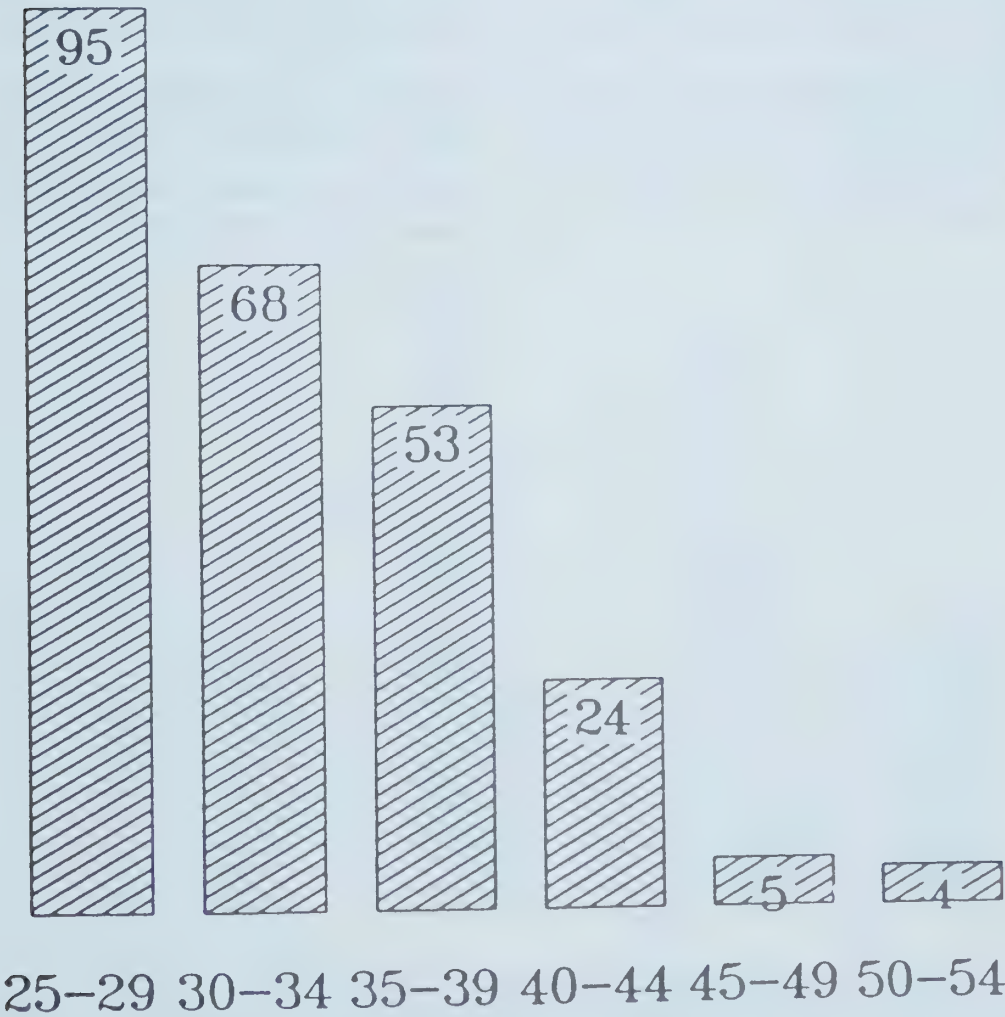


Figure 1: Age profile of adult writers (numbers are the actual numbers of students in each age range)

square analysis of adult question selections among the four questions on each exam, we found that on five of the 10 dates their selection among the four could not be attributed to random distribution (four dates significant at the 1% level, one at the 5% level). These five exam sittings accounted for 60% of the adult writers and were in fact the five largest of the 10 exam sittings.

For these five target dates we contrasted the adult pattern of question choice with that of the youth (writers 20 years and under) by recording the results both numerically and graphically. We also analyzed the relationship between popularity of a question and the success of those choosing it (success determined by pass rate) for both the adult and youth samples by performing a linear least squares analysis as well as higher order correlations. Again, the young adult population was not included in these analyses; their numbers were too small to provide reliable statistics.

Results

The profile of performance by grade received indicates that over the year the adult writers achieved a pass rate of 73.9% compared with the youth pass rate of 61.8%. We illustrate performance by grades in Figure 2 (for each grade category the left-hand bar represents the percentage of adults achieving the grade and the right-hand bar represents the percentage of youth).

With the knowledge that the adult writers were markedly more successful than the youth on these exams, we turned our attention to their topic choices. We determined that question selections were fairly evenly distributed among the four questions of the sets over the year, both for adult

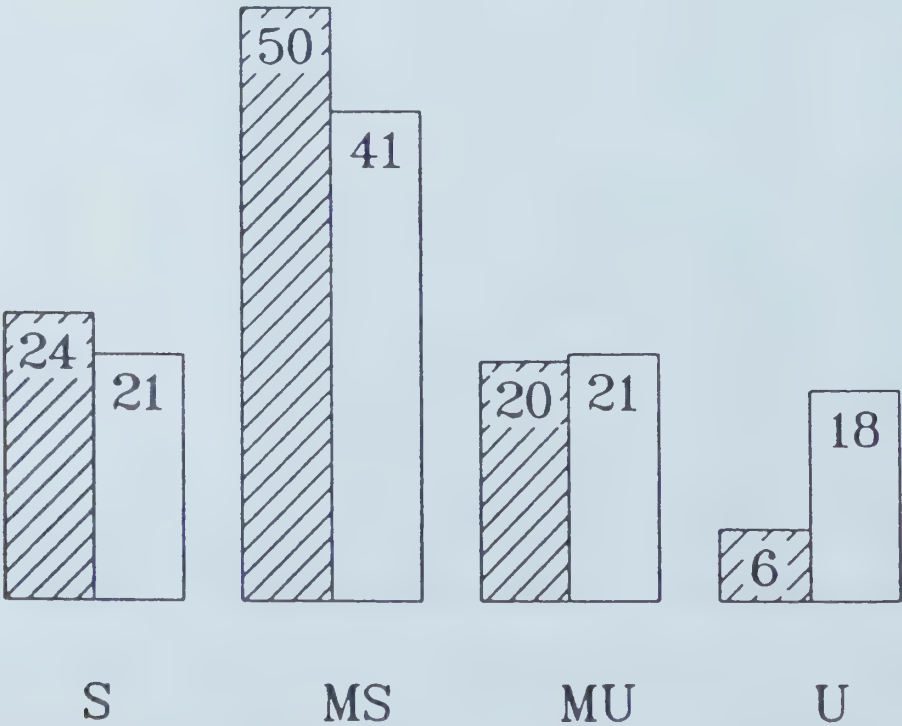


Figure 2: Grade frequencies for adult and youth writers over a one-year period (bars represent % achieving a mark)

writers and for youth; that is, writers did not particularly favor questions appearing first, for example, in a set. From the target dates we had 20 topics to consider (see Appendix 1 for the actual questions) and a sample of writers aged under 21 with which to contrast the adults. The figures below illustrate contrastive topic selection on these five exams. Each bar represents the percentage of students who chose a particular question; each set of four bars totals 100%.

From these figures it is apparent that some topics were popular with both adults and youth, and there were a few which were universally unpopular. But there were also topics which had different appeal for the two age groups.

Three topics had obvious attraction for both age groups: seat belt laws, the value of fashion, and the choice of a city in which to live. These three were among the top five choices on both lists selected by more than 40% of both adults and youth. The two groups also agreed on four of the bottom five choices, topics chosen by fewer than 10% of those writing. The rejected four were freedom of speech at universities, international terrorism, students' union issues, and world trust and understanding. However, in other cases the two groups disagreed sharply on the attractiveness of a topic.

Some topics appealed much more to adults than to youth: the characteristics of an ideal marriage partner, the powers of unions, government support for family farms, and modern improvements in daily life. And the young writers responded to some topics which did not interest adults: a

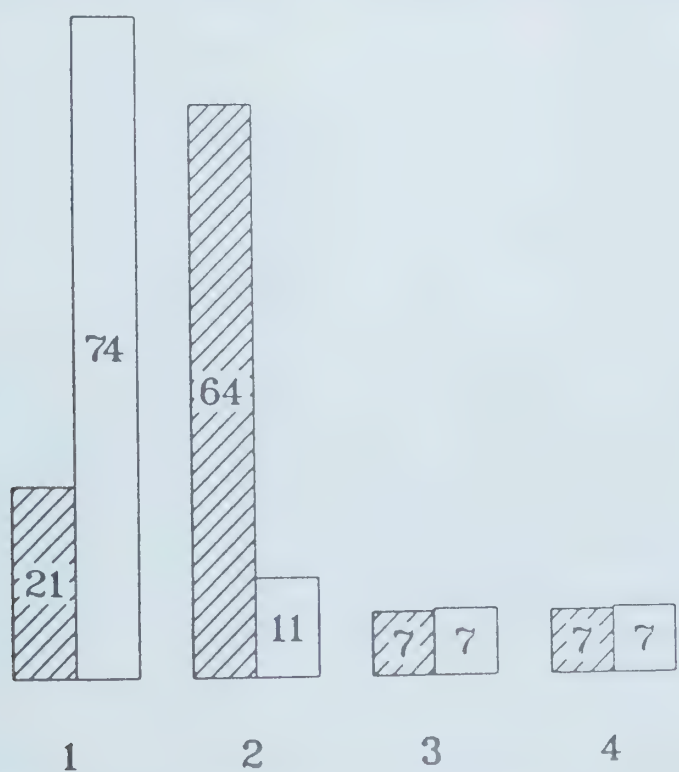


Figure 3: July 24, 1986: Adult and youth question selection
Relative popularity of individual questions with adult (hatched) and youth (clear) writers (bars represent % choosing a question)
Topics: 1. favorite sport 2. ideal marriage partner 3. freedom of speech 4. terrorism



Figure 4: August 6, 1986: Adult and youth question selection
Relative popularity of individual questions with adult (hatched) and youth (clear) writers (bars represent % choosing a question)
Topics: 1. today's youth 2. winning in sports 3. trust among nations 4. 20th-century daily life

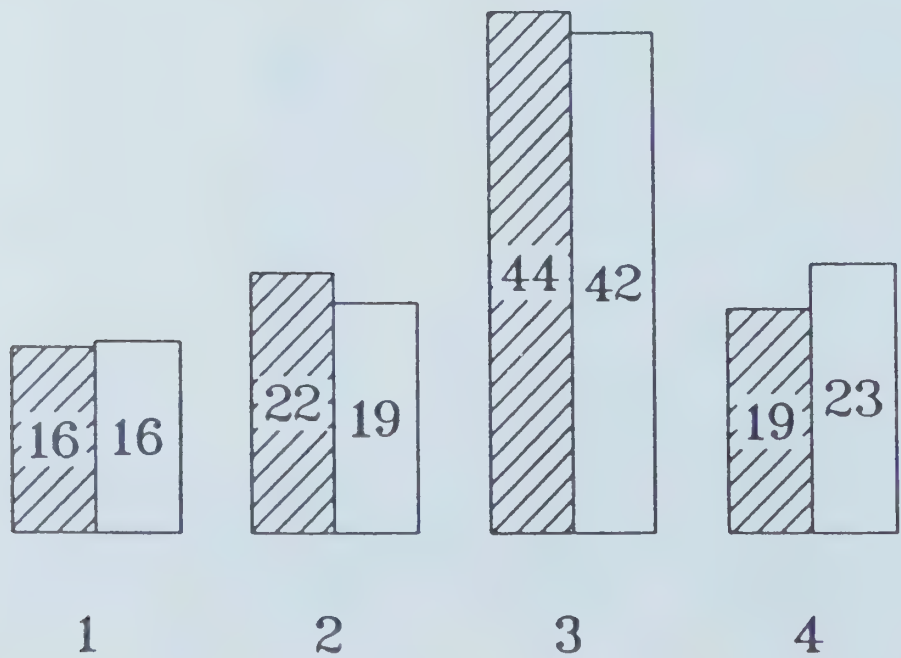


Figure 5: September 5, 1986: Adult and youth question selection
Relative popularity of individual questions with adult (hatched) and youth (clear) writers (bars represent % choosing a question)
Topics: 1. a second language to study 2. education 3. seat belt laws 4. time capsule

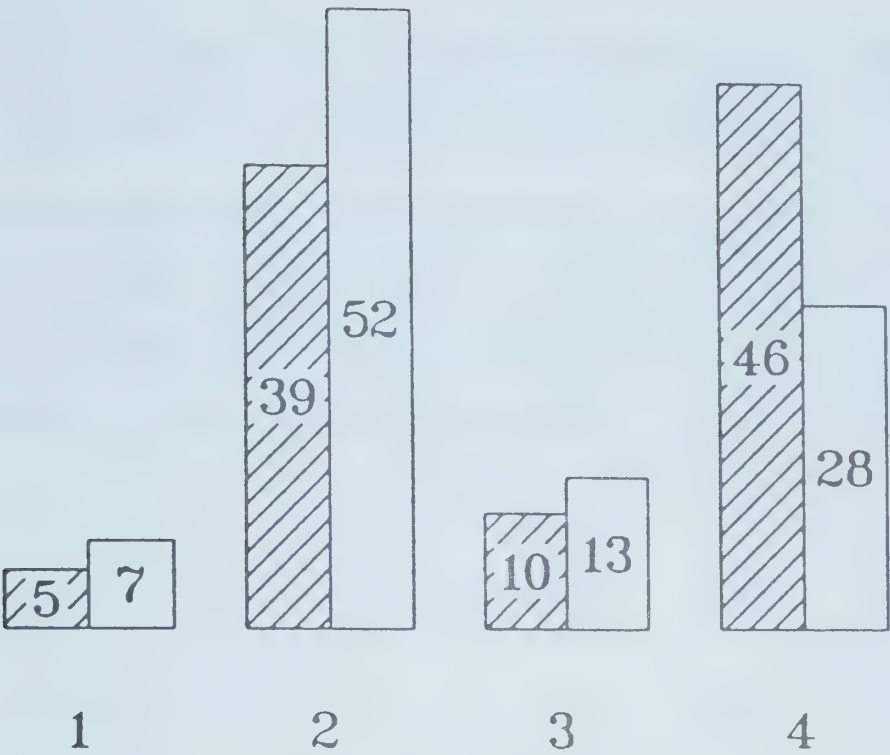


Figure 6: November 29, 1986: Adult and youth question selection
Relative popularity of individual questions with adult (hatched) and youth (clear) writers (bars represent % choosing a question)
Topics: 1. students' union issues 2. fashion 3. humor 4. labor unions

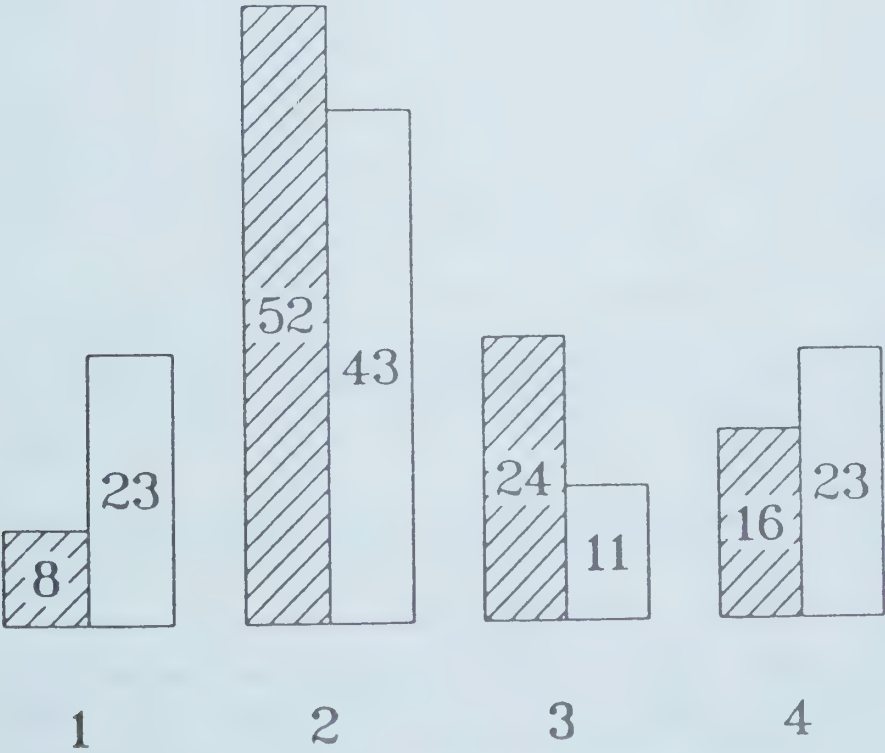


Figure 7: April 6, 1986: Adult and youth question selection
Relative popularity of individual questions with adult (hatched) and youth (clear) writers (bars represent % choosing a question)
Topics: 1. university instructors 2. city to live in 3. family farms 4. what work teaches us

favorite sport or game, the importance of winning in sports, and the teaching abilities of university instructors. But the popularity of a question and the success (as measured by passing the exam) of those who chose it were not correlated.

The linear least squares analysis to determine the relationship between popularity of a question and the success of the students choosing it revealed no correlation whatsoever, either for adults (Figure 8) or for youth (Figure 9). The scatter diagrams in the figures illustrate this graphically. The slope of the regression line for adults was .073, for youth .056; clearly, neither adults nor youth in this sample fared any better or worse on unpopular or

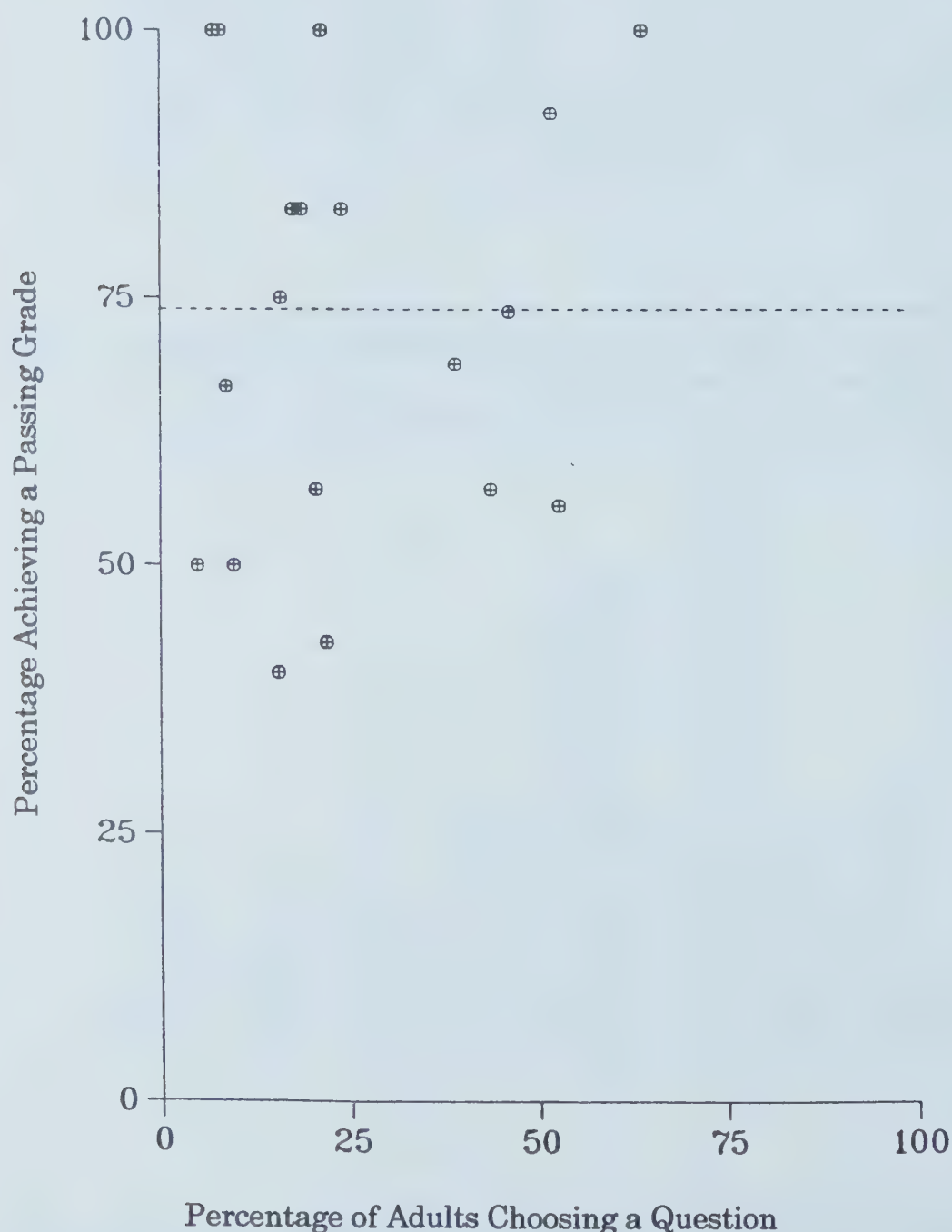


Figure 8: A scatter diagram relating % achieving a passing grade to the % choosing a given question for adult writers (the dashed line represents the baseline pass rate for this group)

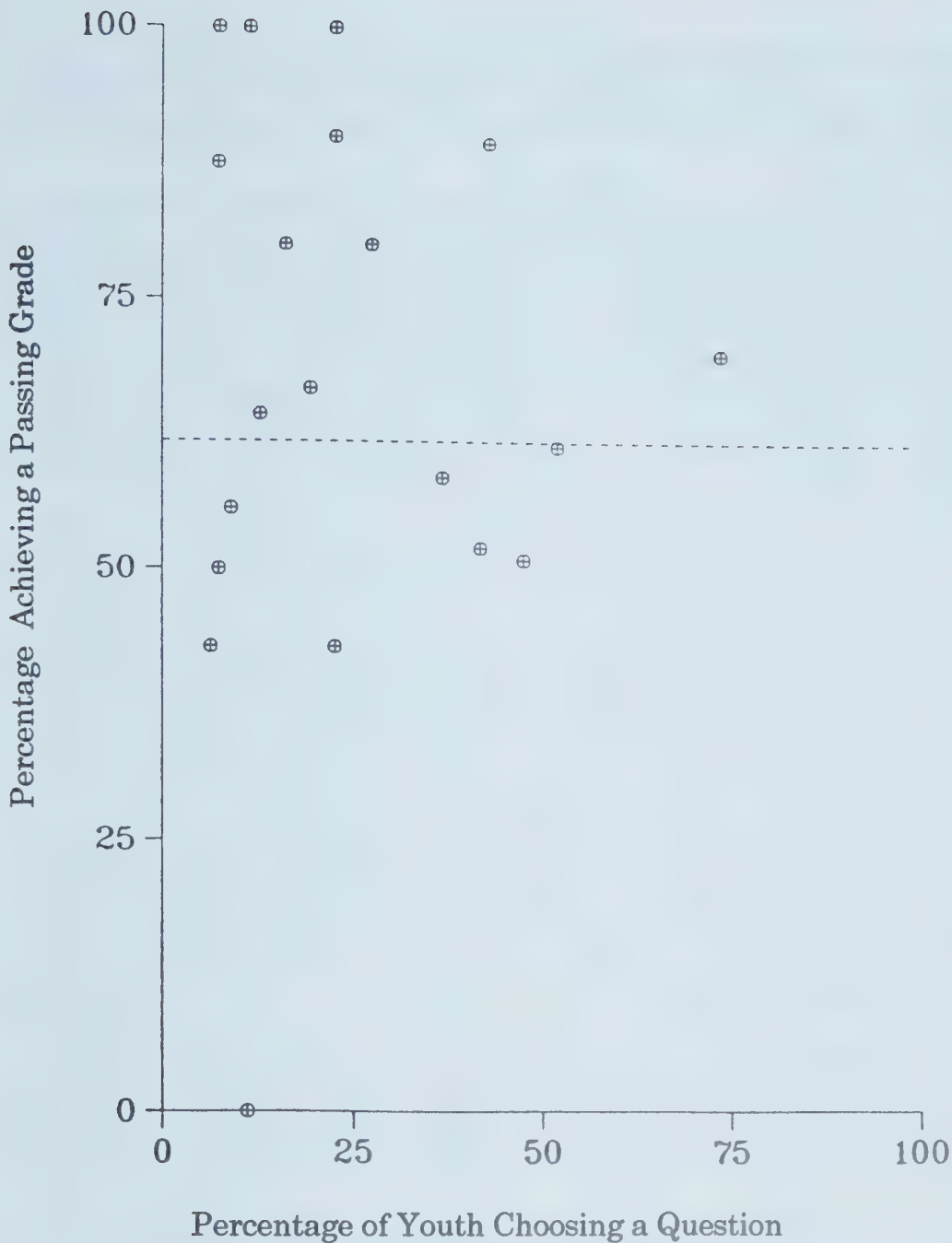


Figure 9: A scatter diagram relating % achieving a passing grade to the % choosing a given question for youth writers (the dashed line represents the baseline pass rate for this group)

popular questions. Three higher order correlations were also tested unsuccessfully.

Discussion

Most questions on the AUWCT are closest to the “moderate information load” described by Brossell (1983) in his discussion of rhetorical specification in essay examination topics. The AUWCT questions are more than “briefly stated subjects” yet have less than “a full rhetorical context.” They may elicit either “reflexive” or “extensive” responses (Emig, 1971) and occasionally impassioned direct address. The official guidelines require that questions be of general interest, raise only one issue, and avoid highly emo-

tional issues. On each paper, of four questions, three are expected to be answerable by non-Albertans and at least one to be of particular interest to students from rural backgrounds, from nonacademic backgrounds, and from outside Canada.

The actual questions are painstakingly written, evaluated, discussed and selected—months in advance—by individual instructors, departments at the four provincial universities, and the interuniversity committee which oversees the exam and makes the final topic selections. Still, students (and sometimes markers) commonly complain that they “didn’t like any of the questions.” Students also assert that they “didn’t know anything about any of the topics,” which no doubt is sometimes true. The instructions on the exam disclaim the need for any “specific knowledge” of a topic, but that may be of little comfort to the exam writers.

It is apparent that in this sample, adult writers performed more successfully than youth in their responses. These adults are not the “basic writers” often discussed in adult composition research, and clearly it would be a mistake to assume that adult students’ writing will be less proficient than that of younger students. But these older writers differ from their juniors not only in performance but also in topic selections, which we can only assume is deliberate and reflects to some extent personal interests and subject preferences. What, then, can we infer from their topic choices?

First, we note a tendency on the part of the adult writers to choose topics on which they can bring to bear rather broad experience in life and work: the powers of labor unions, the qualities of an ideal marriage partner, and assessment of today’s youth. Youth, by contrast, tended to favor topics of more immediate and less philosophical interest: sport, fashion, or seat belt legislation. Shook (1985) advocates that we avoid setting personal, introspective topics for adults. But the adults in our study by no means shunned such topics; in fact they seem to have sought them out, given that they chose from a set of four topics which could elicit either reflexive or extensive responses.

In selecting a particular topic, we suggest that these adults (and these youth) chose topics about which they felt authorized or qualified to write. Hoetker and Brossell (1986) refer to “content-fair” topics as “sets of topics that provide the great majority of examinees with the chance to write on a subject about which they know a good deal” (p. 329). All the topics used in the AUWCT evaluations seem in this sense content-fair, but apparently the older writers felt better qualified to write on certain topics than on others. Those that they chose seem to draw less on knowledge about current events or topical issues than on experience of life, and it is here, we suspect, that these adults found their sense of authority to write.

It is interesting to note that for neither group was performance on a topic correlated with its popularity, thus confirming the results of Brossell (1986) and contrasting with the findings of Gee (1985). Gee, finding the frequency of selection of a topic inversely proportional to the score received, suggests that markers grow bored and lethargic marking the same topic over and over, then “perk up for those rare few on a different topic” (p. 84)

and that marked interest or lack of it may result in higher or lower grades. However, in this case it may be that what is a majority choice for adults is a minority choice overall and perhaps of more appeal to jaded markers.

Conclusions

The special needs, interests, and abilities of the adult students who constitute an increasingly large proportion of university and college student populations require the consideration of these institutions. This study and others indicate that many adult students, although often formally unprepared, bring to their studies high levels of readiness, maturity, and communication skills that make them in many ways “advantaged” students.

This study has several implications for those teaching composition to adult students or designing instruments of writing assessment. First, it seems clear that one need not choose topics targeted at a particular age group’s experience and interests, despite the tendency of the adult writers to select topics in line with their more extensive life experiences; the adult writers in this study chose a broad range of topics, adapting their responses to their own experience and range of knowledge, and succeeded very well. Second, one need not avoid setting topics which are personal or introspective so long as the writers are not compelled to choose them. Given a choice, our adult writers often chose to write with the authority of personal experience and sometimes with what we sensed to be a pleasant introspection, a vein of reflection generally missing in the papers of the younger writers. Finally we can only conclude that, in general, age is no disadvantage to these adults as writers; markers often comment appreciatively on the “mature voice” behind adult writing, however unprepared or unpracticed the writers.

Research concerning writing topics has focused primarily on topic format and effect on writing performance. We suggest that a shift in attention be made from topic format to the writer. What factors make a topic both accessible and appealing to writers of different ages? A greater focus on writers’ perceptions of a topic and motivations for choosing it would be a useful addition to current research in topic analysis.

Finally, there seems to be a gap in composition research in the area of adult writing skills, with the exception of adult basic education. Adults now appear in all our university, college, and community composition courses and workshops; we need to know more about how they write, why they write as they do, and how best to help them improve their writing. For years composition research has concentrated largely on the writing of children and youth; attention to the writing of adults is long overdue.

Acknowledgments

The authors wish to thank the English Language Proficiency Committee of the University of Lethbridge for its cooperation, Peter Haney of the Institutional Research Office for his assistance, and Seamus O’Shea of the University of Lethbridge for his generous help with statistics and graphics.

References

- Brand, A.G., & House, G. (1987). Relationships between types of assignments, writer variables, and emotions of college composition students. *Alberta Journal of Educational Research*, 33, 21-32.
- Brossell, G. (1983). Rhetorical specification in essay examination topics. *College English*, 45, 165-173.
- Brossell, G. (1986). Current research and unanswered questions in writing assessment. In K.L. Greenberg, H.S. Wiener, & R.A. Donovan (Eds.), *Writing assessment: Issues and strategies* (pp. 168-182). New York: Longman.
- Brossell, G., & Ash, B.H. (1984). An experiment with the wording of essay topics. *College Composition and Communication*, 35, 423-425.
- Campbell, D.D. (1984). *The new majority: Adult learners in the university*. Edmonton: University of Alberta Press.
- Carlman, N. (1986). Topic differences on writing tests: How much do they matter? *English Quarterly*, 19, 39-49.
- Connors, P. (1982). Some attitudes of returning or older students of composition. *College Composition and Communication*, 33, 263-266.
- Emig, J. (1971). *The composing process of twelfth graders*. Urbana, IL: National Council of Teachers of English.
- Gee, T.W. (1985). Of ships and sealing wax: Drafting and revising processes in grade twelve students' examination writing. *English Quarterly*, 18, 82-88.
- Gee, T.W. (1987). Drafting and revising processes in grade twelve students' examination writing. *Alberta Journal of Educational Research*, 23, 96-114.
- Hayes, J.R., & Flower, L.S. (1986). Writing research and the writer. *American Psychologist*, 41, 1106-1113.
- Hoetker, J., (1982). Essay examination topics and students' writing. *College Composition and Communication*, 33, 377-392.
- Hoetker, J., & Brossell, G. (1986). A procedure for writing content-fair essay examination topics for large-scale writing assessments. *College Composition and Communication*, 37, 328-335.
- Langer, J. (1983). *Effects of topic knowledge on the quality and coherence of informational writing*. ERIC Document Reproduction Service No. ED 234 418.
- Nagy, P., Evans, P.J.A., & Robinson, F.G. (in press). Exploratory analysis of disagreement among holistic essay scores. *Alberta Journal of Educational Research*.
- Odell, L., Cooper, C.R., & Courts, C. (1978). Discourse theory: Implications for research in composing. In C.R. Cooper & L. Odell (Eds.), *Research in composing: Points of departure* (pp. 1-12). Urbana, IL: National Council of Teachers of English.
- Shook, R. (1985). Let's not write about ourselves. *Teaching in the Two-Year College*, 12, 203-206.
- University of Lethbridge. (1973). *The University of Lethbridge Statistics 1972-73*. Lethbridge, AB: Office of the Registrar.
- University of Lethbridge. (1988). *The University of Lethbridge Facts Book 1987-88*. Advance Information Institutional Research Office, Lethbridge, AB.

Appendix 1: AUWCT Questions from Target Dates

July 24, 1986

1. What is your favorite sport or game, and what makes it enjoyable?
2. What are the characteristics of the ideal marriage partner? Support your conclusions.
3. Recently, South African Ambassador Glen Babb was prohibited from engaging in a debate at the University of Toronto Law School. Opponents of his visit argued that Babb would be defending an indefensible policy: apartheid. Other students argued that the refusal to allow Babb to participate in a public debate was a violation of a basic human right: freedom of speech.

What restrictions, if any, should universities place on the selection of speakers for open debate? Support your conclusions.

4. What steps do you think should be taken to help reduce the threat of terrorism in the world today? Support your conclusions.

August 6, 1986

1. Some older people say that young people today are generally self-centered and greedy. They argue that young people don't care about the welfare of others or about causes that don't concern them; they want only to obtain high-paying jobs and a luxurious way of life.
Do you agree or disagree with this assessment of today's youth? Support your conclusions.
2. "Winning isn't everything. It's the only thing."—Former football coach Vince Lombardi
Do you think that winning is the primary objective in sports? Support your conclusions.
3. Many people who are campaigning for nuclear disarmament argue that ways of promoting understanding and trust among nations must first be found; only then will disarmament be possible.
By what specific means could we promote understanding and trust among nations? Support your conclusions.
4. What person, policy, or invention do you think has made a significant improvement in our daily lives in this century? Support your conclusions.

September 5, 1986

1. If you could choose one language other than your own to study, which would you choose and why? Support your conclusions.
2. What are the qualities of a well-educated person? Support your conclusions.
3. Some people argue that laws which make the use of seat belts compulsory curtail the freedom of the individual; others, however, maintain that such laws are necessary.
Should the wearing of seat belts in automobiles be compulsory? Support your conclusions.
4. The time capsule is a wonderful notion. The idea is to collect contemporary artifacts in a capsule and bury it so that future generations—or perhaps visitors from outer space—will know something about the way we live. —adapted from Richard Cohen, *The Washington Post*
What item or items would you put in a time capsule to characterize your own country as it is in 1986? Support your conclusions.

November 29, 1986

1. If you were President of the Students' Union at your university, what issue or issues would you devote your time to? Support your conclusions.
2. Most of us spend far too much time and money keeping up with fashions. Why do we go to all this trouble? We should do away with

fashions in dress and design an inexpensive, comfortable garment for everyone, men and women, to wear. —Thetis Oliver

Do you agree or disagree with this suggestion? Support your conclusions.

3. A type of humour prevalent in our society is based on stereotyping, the attributing of characteristics to groups of people on the basis of race, sex, or nationality. For example, ethnic jokes and comic strips featuring “typical housewives” are both very popular. Although some people see no harm in this type of humour and may even see it as healthy, others think that it is offensive and that it reinforces harmful prejudices.

Do you see any harm in this type of humor? Support your conclusions.

4. Most people agree that when the labour movement first began, powerful unions were needed to protect workers against appalling conditions and exploitation that verged on slavery. However, some say that today's unions are too powerful: they argue, for example, that unions strike far too often, sometimes even to the detriment of the union workers themselves.

Do you think that the powers of today's unions should be limited? Support your conclusions.

April 6, 1987

1. Although few university students complain about their instructors' knowledge, they often complain about their instructors' abilities to impart that knowledge. Some have suggested that the problem would be solved if all university instructors were trained as teachers. Should all university instructors be trained as teachers? Support your conclusions.
2. If you had the opportunity to live in any city in the world, which would you choose and why? Support your conclusions.
3. Because of the difficulties of the last few years, many small family farms are on the brink of financial disaster, and many people believe that the government should assist these operations. However, farming is becoming more complex and expensive, and it seems likely that soon only the large corporate-owned farm (agribusiness) will be able to profit. Does the family farm deserve support from the government? Support your conclusions.
4. “I don't like work—no man does—but I like what is in work—the chance to find yourself.” —J. Conrad
What can work teach us about ourselves? Support your conclusions.

ANDREW GARROD

Dartmouth College

Promoting Moral Development Through a High School English Curriculum

This study examines the effects of a senior high school English curriculum in stimulating moral reasoning development as defined by Lawrence Kohlberg. The specific hypothesis tested in a pretest/posttest control group design involving 44 grade 10 senior high school students was that an advanced level English literature curriculum focusing on moral issues would promote moral stage change. The major finding of the study was that the curriculum proved effective to a significant degree in stimulating the growth of moral reasoning in preconventional students, but when the performance of the experimental group was compared with that of the control group, there was no overall statistical significance. The discussion section recommends psychological goals, beyond moral reasoning development as defined by Kohlberg, as appropriate for an English curriculum.

Introduction

Much research evidence supports the hypothesis that the developmental level of moral reasoning, as defined by Kohlberg's (1981) cognitive-developmental theory of moral development, can be advanced through group discussion of moral dilemmas and through exposure to reasoning at a higher stage. Academic subjects provide ample opportunity for discussion of ethical and social issues, yet there appears to be no research published documenting the impact of an English literature curriculum which has as an explicit goal the raising of moral reasoning level. This paper first offers a brief review of the literature on classroom moral stage change, and second it describes the effects on moral stage of a pilot curriculum in English literature among grade 10 academic students in New Brunswick.

The investigator's hypothesis was that the use of literary material of high caliber and a pedagogy that emphasizes discussion, exploration, and debate of issues in the lives of literary characters, characters in film, and in the students' own lives—all in a supportive environment—will stimulate the moral reasoning capabilities of the students as they are encouraged to confront the greater adequacy of moral reasoning at a stage above their own.

The investigator recognizes that Kohlberg's theory is cognitive in its emphasis and that there is more to moral development than his ascending stages in conceptions of justice. Nevertheless, Kohlberg offers the re-

Andrew Garrod is an assistant professor of education in the Department of Education at Dartmouth College. His research interests are adolescence, moral development, and moral education.

searcher a well-tested and validated instrument for the assessment of a critical element in moral development.

Review of the Literature

Moral judgment development, according to Kohlberg (1984), proceeds in a step-like manner. Each step or structure in his moral reasoning scale is a qualitative reorganization of a person's pattern of thought that integrates and subsumes the insights achieved at lower steps; each stage represents a more adequate and logical way of understanding the moral world than the previous one. The stages are divided into three levels, each containing two stages. At the *preconventional* level (Stages 1 and 2), characteristic of children and developmentally slow adolescents and adults, the individual is

responsive to cultural rules and labels of good or bad, right or wrong, but interprets these labels either in terms of the physical or the hedonistic consequences of action (punishment, reward, exchange of favors) or in terms of the physical power of those who enunciate the rules and labels. (Scharf, 1978, p. 308)

At the *conventional* level (Stages 3 and 4), the level of the average adolescent and adult, the individual understands and tries to uphold the values of his or her society; the expectations of others—peers and authorities—have been internalized. At the *postconventional* level (Stages 5 and 6), characteristic of very few adults, “customs and social rules are critically examined in terms of universal human rights and duties and universal moral principles” (Jennings, Kilkenney, & Kohlberg, 1983, p. 282). Table 1 clarifies at each stage what is considered right.

In the last 15 years, educational research influenced by Kohlberg's theory of moral development has followed two main lines: (a) interventions in classrooms having the explicit goal of raising the stage of moral reasoning among members of the experimental group, and (b) the creation of democratic high schools, where the researcher has looked specifically at the interaction of the stage of moral atmosphere, the norms of the group, and the individual's structures of moral reasoning. The first line is indebted to Piaget and Inhelder's (1969) and Kohlberg's (1981) interactive models of cognitive and moral development; the second line is more sociological, examining the impact of shared norms on individual moral growth. This paper considers the research on classroom intervention.

There have been many studies to validate the effect of classroom moral discussions on moral reasoning growth. Kohlberg's hypothetical moral dilemmas, originally a tool of assessment, were first used by Blatt and Kohlberg (1975) as a “curriculum” in an attempt to demonstrate that significant growth in moral reasoning, as measured by the Moral Judgment Interview (MJI), could be promoted in a group exposed to discussion of moral dilemmas, in contrast to a control group not similarly exposed. This study and others by researchers such as Schaffer (1974) served as pilot projects for this form of moral education. Typically, a moral dilemma is read aloud by the teacher and followed by class discussion. A mix of students at different stages of moral reasoning has been found preferable because a

Table 1
Stages of Moral Development According to Kohlberg (1981)

<i>Stage</i>	<i>What is considered to be right</i>
Stage 1: Obedience and punishment orientation	To avoid breaking rules backed by punishment, obedience for its own sake, avoiding physical damage to persons and property.
Stage 2: Instrumental purpose and exchange	Following rules only when it is to someone's immediate personal interest; acting to meet one's own interests and letting others do the same; right is an equal exchange, a good deal.
Stage 3: Interpersonal accord and conformity	Living up to what is expected by people close to you or what people generally expect of those in your role; being good is important.
Stage 4: Social accord and system maintenance	Fulfilling the actual duties to which you have agreed; laws are always to be upheld except in extreme cases where they conflict with other fixed social duties; right is also contributing to society, the group, or institution.
Stage 5: Social contract, utility, individual rights	Being aware that people hold a variety of values and opinions, that most values and rules are relative to your group but should usually be upheld because they are the social contract; some nonrelative values and rights like life and liberty, however, must be upheld in any society regardless of the majority opinion.
Stage 6: Universal ethical principles	Following self-chosen ethical principles; particular laws or social agreements are usually valid because they rest on such principles; when laws violate these principles, one acts in accordance with the principle; principles are universal principles of justice; the equality of human rights and respect for the dignity of human beings as individual persons; the reason for doing right is the belief, as a rational person, in the validity of universal moral principles and a sense of personal commitment to them.

Note: Stages 5 and 6 are not distinguished for research purposes; there is also a transition stage (e.g., 2/3, 3/4) between each of the stages.

variety of viewpoints adds to students' cognitive conflict and thus is instrumental to growth. It is assumed that the teacher's answers will reflect a higher stage of moral reasoning and serve as a model for the most mature students in the group, thereby promoting growth.

Most subsequent curricula inspired by Kohlberg's theory have used a similar Socratic pedagogy. In his summary of research in values clarifica-

tion and moral development curricula, Lockwood (1978) analyzes 11 moral education studies, dividing them between those that employ direct moral discussion and those that employ psychological activities (e.g., peer counseling). The direct discussion approach generally produces significant development in moral reasoning. According to Lockwood, the best-designed studies showed a mean advance of almost one half stage over a semester or an academic year, as opposed to one-third stage advances achieved by the psychological awareness approach. All subjects, of course, did not advance consistently; some students developed, while others remained at the same stage. Lockwood explains that the mean developmental increases are in the Stage 2 to Stage 3 range.

In the most recent review of moral education programs and curricular effects on participants' moral reasoning, Higgins (1980) finds an increase between one third and one half stage in experimental groups. Both Higgins and Lockwood confirm that in interventions where psychological activities predominate (as opposed to moral discussion), moral growth, as measured by the MJT, is modest.

In sum, then, there have been studies that used hypothetical dilemmas exclusively as the curriculum for moral education and others that have attempted to weave into the discussions activities that relate more generally to psychological growth—for example, peer counseling, cross-age teaching. Few have attempted to integrate moral education into an academic discipline, and the results of such studies have been contradictory and inconclusive. Lieberman (1979) looked at the effects on moral growth of a junior high school 10-week unit, *Facing History and Ourselves: The Holocaust and Human Behavior*, and found that the experimental group did not increase in moral judgment significantly more than the control group, although significant increases in interpersonal awareness stage were found for the experimental group. Similarly, in two American history classes taught by Ladenburg (1977) and his colleague Oakes at a Massachusetts high school—where role-playing of historical situations and engagement in hypothetical direct moral discussion were featured—neither group showed any moral growth.

In contrast, a major study of secondary school moral discussion programs led by social studies teachers in the Boston and Pittsburgh areas as reported by Colby, Kohlberg, Fenton, Speicher-Dubin, and Lieberman (1977) found significant increases in moral judgment reasoning in comparison with controls. This study provides a research base for hypothesizing that stage change can be realized through moral discussion within an academic discipline and through the collective role-playing that is a part of the study of fiction and drama in an English class that relies on the Socratic method. Additional support for hypothesizing stage change as a result of the investigator's curriculum, with its blend of role-playing, discussion of moral and social issues in literary texts, theater games and drama, comes from the year-long course taught by Sullivan (1975) to juniors and seniors in high school as a social studies elective. His course included such elements as

moral psychology and philosophy, moral discussions, and counseling and empathy training, and it achieved a striking mean increase of one half stage.

Teachers of literature are in an especially favorable position for assuming the responsibilities of moral education because so many poets and writers address themselves to moral issues. Sometimes justice is the focal point of the work, as in Lee's *To Kill a Mockingbird* or in Lawrence and Lee's *Inherit the Wind*; sometimes the moral issues are more peripheral, as in *Huckleberry Finn* or *A Separate Peace*. Each of these works is charged with moral problems which, through discussion, can promote the moral growth of young people.

In creating this experimental curriculum in moral development and literature (Garrod & Bramble, 1977), Kohlberg's (1981) cognitive-developmental approach to moral education was used because it is a conceptual tool which, in the face of conflicting values and moral systems, helps individuals to determine the fairest and most adequate course of action. Furthermore, Kohlberg's theory can be used successfully by teachers to stimulate moral development without recourse to the "preaching" of moral content (which many educators regard as objectionable).

The curriculum under study has the potential for initiating the process of examination and evaluation that leads to a careful reconsideration of inherited values by the young. Whether those values are reaffirmed or rejected, the student will have ceased to be a person whose morality is unquestioned. In this curriculum, literature is the vehicle for critical thinking and moral development. The following discussions of classroom process show how critical thinking and moral reasoning development were encouraged in this curriculum.

Classroom Process

Blatt (1970) hypothesized that the cognitive-developmental approach to moral education might have practical as well as theoretical applications. He discovered that "moral discussion" was a powerful method of stimulating the development of moral reasoning. As a result, moral discussion has recently been introduced in many junior and senior high school social studies classrooms. However, the English curriculum used here takes a different approach. Whereas the social studies curricula based on the cognitive-developmental mode have relied heavily on case studies and hypothetical dilemmas, in this investigator's curriculum, literature is the source of discussion material. As in the social studies courses, moral discussion is a key strategy for stimulating moral development. The characters and situations in this curriculum's texts exhibit clearly delineated values and attitudes. According to their established values, these fictional characters make decisions whose consequences they must endure for good or for ill.

When the quality of the literature is high, fictional characters manifest points of view as valid as those held by members of the discussion group. A significant degree of role-taking (the ability to take on another's social

perspective) is required if students are to appreciate these fictional perspectives. Students must also learn to listen carefully to one another and try to understand points of view they may regard as antagonistic to, or exclusive of, their own. Students were encouraged to show as much consideration for the characters in the selected works as they would extend to their peers in the discussion group.

Normally, student role-taking in a moral discussion is confined to perspectives held by members of the immediate group, but this curriculum invited students to assume the viewpoint of fictional characters as well. This role-taking was generated by posing questions designed to elicit *value identification* as well as moral reasoning. Value identification is the role-taking process by which students attempt to *identify* the specific attitudes and values they perceive to be held by others. When a member of a discussion group is the object of this process, verification can always be obtained through feedback. When the process involves a fictional character, however, direct feedback is not possible, and discussion participants must identify the character's values by a collective role-taking process in which students compare their impressions of the fictional character with those of their peers. Applying this process to the values and attitudes of fictional characters (which many experienced English teachers do as a matter of course) may impose stringent demands on the role-taking powers of discussion participants, but only in this way can characters be brought to life for students.

Curriculum Implementation

In practice, students in this study explored the moral problem as embodied in the text before moving on to a broader discussion. The text anchored the debate in the concrete and prevented it initially from becoming too hypothetical and abstract.

Questions were phrased to permit responses at various stage levels. Thus, while the content may be similar in the responses of two different students (e.g., that one shouldn't lie), the reasoning behind the response may indicate that the students are at different stages of moral development, agreeing on a particular choice but not on the reasons for making that choice. Consequently, students' responses were carefully probed until their reasoning became clear to all members of the group.

In some texts, the exact nature of the moral problem may be debatable, as in *A Separate Peace*, where the psychology of the characters is complex and the author's interest is in psychology as well as justice. The developmental approach to moral education stimulates cognitive and moral dissonance as a means of promoting moral development rather than focusing on the "correct" identification of the moral problem. It is perhaps unnecessary that every moral issue be sharply defined; it is enough that students find sufficient substance to trigger an interaction between their own moral viewpoint and that of the fictional character under consideration. Through interaction—between student and literature, student and student, or student and teacher—moral development can occur.

Huckleberry Finn is an appropriate text to illustrate how a moral discussion approach can inform students' reading and appreciation of a novel; it was selected to apply Kohlberg's model to literature because Twain's book is rich in themes of friendship. As the novel begins, Huck's moral development has been temporarily checked by the smothering effects of Pap's bullying, Tom's domination, and the Widow's motherly reforming. But Jim seems to intuit that caring for human beings often necessitates letting go of them—that showing respect for the dignity of the individual means honoring the right to make autonomous decisions. Of all the characters in the book, only Jim shows Huck this total acceptance. But freedom implies responsibility, and in the egalitarian environment of the raft and the river, Huck must bear the burden of, and the accountability for, his own moral judgments. No external authority can rescue him from his moral predicaments; he must decide for himself what is right, and it is this freedom to agonize over moral dilemmas that leads to Huck's growth.

The incidents occurring between chapters 16 and 31 are a constant source of cognitive and moral conflict for Huck, and the result seems to be a transition, in terms of Kohlberg's model, from the preconventional to the conventional level of moral reasoning. In these chapters, Huck's continuing dilemma is conflicting personal loyalties—to Miss Watson on the one hand and to Jim on the other. These loyalties suggest a Stage 3 orientation, and until the climactic episode at the end of chapter 33, Huck's moral reasoning vacillates between Stage 2 and Stage 3. As his relationship with Jim continues to blossom, he grows increasingly more concerned with Jim's welfare at the expense of his own. This shift reflects a rejection of self-interest (Stage 2) as behavioral motivation in favor of trust, respect, loyalty, and gratitude—qualities characteristic of Stage 3 moral reasoning.

A caveat is in order. Obviously, a Moral Judgment Interview cannot be administered to Huck (or to any of the other characters in the selected texts), so unqualified stage claims cannot legitimately be made about the characters. It can nevertheless be argued that the story contains a range of reasoning typical of the transition from Stage 2 to Stage 3. Twain never makes the mistake, for example, of letting Huck argue the evils of slavery or the inalienable rights of humankind—concerns that arise at Stage 5. Without knowledge of stages and structures, he understood human nature well enough to avoid this sort of glaring incongruity and portrayed Huck's level of moral reasoning consistently in the Stage 2-3 range, typical of a boy his age. *The Adventures of Huckleberry Finn* may have stylistic and narrative flaws, but in his young protagonist, Twain offers a true-to-life character for whom there are no shortcuts to moral development.

To demonstrate more clearly the nature of the curriculum, a few sample chapter synopses and study questions from the unit on *Huckleberry Finn* are given in the appendix. The remainder of this article includes a detailed description of the method and results of the investigator's study using this curriculum. In addition, the strengths and weaknesses of this study are con-

sidered, as are suggestions for future research in moral development and the study of English literature.

Method

Subjects

A total of 44 high school students participated in the study, 18 boys and 26 girls. There were a total of 21 students in the experimental group and 23 in the comparison group. Only those who completed both pretest and posttest were included in the analysis. Both groups were grade 10 enriched English classes. Students opt for a grade 10 enriched English class on the advice of their junior high school English teacher; the main requirements for selection are at least average achievement in English studies and the teacher's assessment that the student is highly motivated.

The groups appeared similar in several respects. Variation in SES, IQ scores, and achievement tests were similar for both groups. The sample reflected a wide variety of family backgrounds, ranging from parents who were university lecturers to those who were dockyard workers and to the unemployed. The average age of each group was estimated to be 15.5 years at the start of the academic year and 16.3 years at its termination.

Design

The design was a quasi-experimental one, closest to Campbell and Stanley's (1963) tenth design, "the Non-equivalent Control Group Design." As indicated above, students were not randomly assigned to groups; rather, the two teachers were randomly assigned to in situ classes—the two grade 10 enriched classes. In other words, the assignment of the "treatment" to one group or another was random. The comparison group was taught by a different teacher, and both she and the teacher-researcher used Socratic discussion extensively in their pedagogy. The experimental and comparison groups were administered the Moral Judgment Interview as pretest (Form A) and posttest (Form B), in each case in written form:

	Time 1		Time 2
Experimental Group	0	X	0
Comparison Group	0		0

The Treatment: The Curriculum

The treatment was a year-long English literature curriculum constructed and taught by the investigator. The comparison group curriculum had some features in common with the treatment—for example, both classes studied a Shakespeare comedy and used a similar text of narrative and expository readings. The treatment group, however, studied *A Separate Peace*, *Huckleberry Finn*, *Antigone*, *David*, and *To Kill a Mockingbird*, texts selected for themes appropriate to the concerns of adolescents and for the moral issues they embody.

Discussions occurred both in small groups and with the whole class; discussion questions were addressed individually in homework assignments or taken up collectively after the reading of a text. Variety of approach to discussions was a consistent consideration. Debates, formal and informal, centered on issues and themes in the literature. Written work was assigned at different stages of the study and discussion of a text. Both spontaneous responses and answers informed by group discussion were elicited throughout the year. Several feature and documentary films, integrated with the themes of the curriculum, were viewed and discussed. In addition, the curriculum for both groups included dramatization of scenes from Shakespeare's *Twelfth Night*, theater games, and the staging of one-act plays in which all students performed.

Measures

The Standard Form Moral Judgment Interview and Scoring System (Colby, Gibbs, Kohlberg, Speicher-Dubin, & Candee, 1979) was used for both pre- and posttest assessments. For all the students, Form A was used as a pretest, Form B as a posttest. In light of the proven greater difficulty of Form B, a scale score was provided in the analysis. All students filled out the MJIs during class time. Oral probes were provided as students were administered the in-class written protocols.

The Moral Judgment Interviews were scored by trained raters blind to both treatment group and times of testing. Scores reported are global stage scores indicating major stage usage and Moral Maturity Scores. Moral Maturity Scores (MMS) are weighted averages of stage use across the total interview. They range from 100 (Stage 1) to 500 (Stage 5), with a score of 350, for example, indicating a protocol midway between Stages 3 and 4. The global stage score, true to developmental theory, presents the data in a more qualitative way; the MMS treats the data as a continuous variable.

Results

In keeping with other reports of intervention studies in cognitive developmental moral education, the analysis considers the data as a continuous variable and examines them in a structural way more true to its theoretical orientation. Thus the format for the presentation of results consists of an analysis of the data as a continuous variable followed by a structural analysis of the data. In each analysis, the data are further examined by stage to assess the differential impact of the curriculum.

The major finding of the study was that the curriculum proved effective in stimulating the growth of moral reasoning in preconventional students (those scoring below Stage 3) and that it led to increases in moral judgment maturity comparable to the largest increases obtained in other studies (e.g., Harris, 1976; Sullivan, 1975). However, when the increases in the experimental group are compared with those in the comparison group, there is no overall statistical difference.

Because it was impossible to randomly assign students to treatment and comparison groups, a t-test was performed to test for initial equivalency of

the groups. The t-test indicated that the difference between the two groups on the pretest was statistically significant ($t=3.15$; $p=.003$); the experimental group exhibited more advanced moral reasoning ($x=291.1$) initially than did the comparison group ($x=270.4$).

In an examination of the mean change in MMS from pretest to posttest for experimental and comparison groups, a t-test revealed that the 42.1 MMS increase of the experimental group, when compared to the 32.4 MMS increase of the comparison group, was not statistically significant ($t=2.10$; n.s.).

Further tests were carried out to see if the curriculum had a differential and significant impact on those who scored initially at different stages—that is, at Stage 2 and Stage 2/3, at Stage 3, and at Stage 3/4. Scores indicating stage mixture were not differentiated by major and minor stage. Thus scores of 2(3) and 3(2) were treated as equivalent (2/3), as were 3(4) and 4(3).

In light of the large difference in variance, a separate variance estimate was used (Table 2), and the t-test revealed that the increase of 62.3 MMS of the Stage 2 and 2/3 students in the experimental group was significantly larger than the advance of 39.4 MMS of the comparison group ($t=3.10$; $p=.008$). This finding is particularly striking in that it indicates that the curriculum was highly effective in stimulating the development of moral reasoning in students at Stages 2 and 2/3.

A similar analysis of change scores (Table 3) was made for those from both groups who scored at Stage 3 on the pretest. Despite the statistically insignificant difference of 13.3 MMS in mean increase between those in the experimental group who scored on the pretest at Stage 3 and those in the comparison group who scored at the same level, the direction of the difference was markedly in favor of the experimental group.

An analysis of the results of those scoring initially at Stage 3/4 is not presented in full because the number of students from both groups was small (five) and the increases modest (7.5 MMS gain with the experimental students and 4.4 MMS with the comparison students, respectively). In addition, a two-way analysis of variance was carried out on the data using sex and group as factors, with no significant differences or interactions found.

Table 2
Change Scores for Stages 2 and 2/3 Students

	Number of Cases	Mean Increase	Standard Deviation	t-value (2-tail prob.)
Experimental	5	62.3	6.1	3.10* ($p = .008$)
Comparison	12	39.4	23.7	

*Based on a separate variance estimate.

Table 3
Change Scores for Stage 3 Students

	Number of Cases	Mean Increase	Standard Deviation	t-value (2-tail prob.)
Experimental	13	42.4	27.3	1.11
Comparison	9	29.1	27.5	(n.s.)

In Table 4 the strong upward trend in the experimental group is clearly evident. The left-hand side of Table 4 reveals dramatically that in the experimental group all who started at Stage 2/3 (five students) went up a full stage. Of the 13 students who started the year at Stage 3, only two failed to move upward, and 11 (84.6%) moved up a half-stage. Those already at Stage 3/4, few in number (three), did more poorly—one regressed a full stage, one regressed a half-stage, and one remained stationary. No one reached Stage 4. These findings suggest the powerful effect of the curriculum on students at two stages: It had maximum impact on Stage 2/3 students and a considerable effect on those pretesting at Stage 3. Of all the students in this group, 95.3% finished the treatment at Stages 3 and 3/4. The regressions may be due to scoring error or may be the result of administering the MJI in written form. In such circumstances, given the absence of probing and possible student inertia, Higgins (1980) reports that regression is not unusual. In contrast, the comparison group results (right-hand side of Table 4) indicate that only 78.2% finished their course at Stages 3 and 3/4 and that 27.1% failed to advance beyond Stage 2/3. As in the experimental group, there was no advance for those who scored initially at Stage 3/4.

Table 5 reports patterns of stage change from pretest to posttest for the experimental and comparison groups. It can be seen that 23.8% of the experimental group gained a full stage, as opposed to only 8.7% of the comparison group who did so. However, the movement up a half stage was remarkably similar—52.4% in the experimental group and 52.5% in the control group.

Only two subjects in the experimental group showed a decline in moral stage; in the comparison group, three subjects showed such a decline. In the experimental group, 14.3% remained at the pretest stage as opposed to 26.1% in the comparison group, but changes in the experimental group generally followed the pattern of sequential and step change found in natural development. The analysis revealed no statistically significant differences in patterns of stage change for the two groups.

Completing a structural analysis of the data of those who initially scored at Stages 2 and 2/3, a significant difference appears between comparison

Table 4
Structural Analysis of Pretest Versus Posttest

Experimental				Comparison			
Posttest Stage Score				Posttest Stage Score			
Stage 2/3	0	0	5	Stage 2	1	0	1
	(0.0%)	(0.0%)	(23.8%)		(100.0%)	(0.0%)	(4.3%)
Stage 3	0	2	11	Stage 2/3	2	7	11
	(0.0%)	(15.4%)	(84.6%)		(18.2%)	(63.6%)	(47.8%)
Stage 3/4	1	1	1	Stage 3	2	3	9
	(33.3%)	(33.3%)	(33.3%)		(8.7%)	(13.0%)	(39.1%)
Total	1	3	17	Total	5	11	23
	(4.8%)	(14.3%)	(81.0%)		(21.7%)	(47.8%)	(100.0%)
Pretest Stage Score				Pretest Stage Score			
Stage 2/3	0	0	5	Stage 2	1	0	1
	(0.0%)	(0.0%)	(23.8%)		(100.0%)	(0.0%)	(4.3%)
Stage 3	0	2	11	Stage 2/3	2	7	11
	(0.0%)	(15.4%)	(84.6%)		(18.2%)	(63.6%)	(47.8%)
Stage 3/4	1	1	1	Stage 3	2	3	9
	(33.3%)	(33.3%)	(33.3%)		(8.7%)	(13.0%)	(39.1%)
Total	1	3	17	Total	5	11	23
	(4.8%)	(14.3%)	(81.0%)		(21.7%)	(47.8%)	(100.0%)

Table 5
Structural Analysis of Stage Change:
Experimental Versus Comparison

Group	-1	-1/2	0	+ 1/2	+ 1	Total
Experimental	1 (4.8%)	1 (4.8%)	3 (14.3%)	11 (52.4%)	5 (23.8%)	21 (47.7%)
Comparison	0 (0.0%)	3 (13.0%)	6 (26.1%)	12 (52.2%)	2 (8.7%)	23 (52.3%)
Total	1 (2.3%)	4 (9.1%)	9 (20.5%)	23 (52.3%)	7 (15.9%)	44 (100.0%)

and experimental groups ($X^2_{(s)}=10.12$; $p=.0063$). As exhibited by Table 2 in the analysis of the data as a continuous variable, Table 6 illustrates the statistically significant growth of the Stage 2 and 2/3 students.

Similarly, a structural analysis of those who scored at Stage 3 initially indicated a much greater advance in the experimental group than in the comparison group. While $X^2_{(2)}=4.90$ with a significance level of .08—that is, it is not quite significant—the number of half-stage advances in the treatment group (84.6%) was far greater than the number of half-stage advances in the comparison group (44.4%). Thus the curriculum proved effective for those who started at Stage 3, as it did for those who started at Stage 2/3.

The small number of those scoring on the pretest above Stage 3 renders a structural analysis of the data on these students uninformative.

Discussion

The present study examines the effects on moral reasoning of an English literature curriculum. The major finding was that the curriculum proved highly effective in stimulating the growth of moral reasoning in preconventional (Stage 1-2) subjects. The curriculum also had a considerable impact on those who initially scored at Stage 3. An important finding was that the

Table 6
Stage Changes for Stages 2 and 2/3 Students

	0	+ 1/2	+ 1	Total
Experimental	0 (0.0%)	0 (0.0%)	5 (100.0%)	5 (29.4%)
Comparison	2 (16.7%)	8 (66.7%)	2 (16.7%)	12 (70.6%)
Total	2 (11.8%)	8 (47.1%)	7 (41.2%)	17 (100.0%)

increase in moral reasoning in the experimental group was comparable to the largest increases in other studies (Blatt & Kohlberg, 1975; Dozier, 1974; Harris, 1976; Sullivan, 1975). However, when the increase in the experimental group is compared with that in the comparison group, there is no overall statistically significant difference.

More specifically, in qualitative terms and in stage breakdown, all experimental students who initially began the program at Stage 2 and Stage 2½ advanced a complete stage. The advance of those in the experimental group who started at Stage 3 (85% moved up a half-stage) approached significance. Thus the curriculum appeared to have a marked effect on those who scored at the lower end of the stages represented in the treatment group; it proved ineffective with those already at Stage 3¼. Overall, the experimental group advanced a mean of 42 MMS points in comparison to the gain of 32 MMS points in the comparison group.

Consistent with developmental theory, almost all observable movement was toward the next stage, although 11% of the students regressed. These regressions may be due to administration of the MJI in written form—not the best method because students are more likely to respond cryptically with responses that would benefit from probing. Both raters, however, commented on the fullness of all the protocols in this study and remarked on their consequent confidence in scoring.

The discussion focuses on the nature of the developmental gains and then assesses the extent to which the results fitted the investigator's clinical sense of what was occurring. The full-stage gains in the experimental preconventional students (below Stage 3) is perhaps attributable to the content of the curriculum, with its emphasis on issues of friendship and affiliation—issues most pertinent to adolescents at Stages 2/3 and 3—but not as conducive to moving a Stage 3 student toward a societal perspective, the hallmark of Stage 4. Perhaps, too, the presence of many conventional students in the class discussion, in keeping with the theory of arousing cognitive dissonance by exposure to higher-stage arguments, is instrumental to growth for these preconventional subjects. In addition, the preponderance of girls in the experimental group, with their relational orientation, might make marked gains in Stage 4 reasoning, as Kohlberg has measured it, unlikely (see, for example, Gilligan, 1982). While the discussion questioning moved from the text to exploration of general principles (e.g., *Is hanging ever right?*), it is also possible that moral and social problems enshrined in literature invite a more contextual moral consideration. They demand a careful reflection on motivation and consequences in a specific context and thus may not stimulate abstract philosophic discussion of principle as much as a psychological discussion of relationships and responsibilities.

This investigator did not explore polarized moral options with his students because in literature of quality the moral concern is inevitably more complexly presented. Rather, he scrutinized the moral/social view in a particular literary context, which was subsequently related to students' experience. It is debatable whether such an approach, undertaken by many

excellent English teachers, is likely to take students beyond conventional reasoning. This may help to account for the dearth of Stage 4 subjects which is surely attributable to the chronological age of this study's subjects as well. In sum, chronological age, the absence of concentration on justice issues, and the focus on psychological understanding of motivations and consequences may be less likely to move students toward Stages 4 and 5 of moral reasoning than are the hypothetical moral dilemma discussions of many proven studies (Harris, 1976).

These findings coincide with reviews of moral interventions by Lockwood (1978) and Higgins (1980), who report that the greatest gains are in the Stage 2 and Stage 3 range. Only Harris (1976) was successful in achieving significant increases in which posttest means were above 300 MMS. A consistent finding is the variability of treatment effect on students at different stages. Neither the direct discussion approach nor the psychological approach to moral education has so far proved instrumental in stimulating reasoning beyond Stage 4 with adolescent populations.

The findings in the experimental group did match the researcher's clinical sense of what was occurring. The students were amazingly vociferous and argumentative. An abundant variety of moral reasoning stages was apparent in the class, and students were intensely involved in discussion. The atmosphere in the classroom was marked by increasing openness and a deepening engagement in literature, ideas, and self-exploration. Throughout the course of the year, and without apparent provocation, students commented that they had reassessed their position and reasons from those expressed on their pretest MJJ; they seemed more ready to claim the new stance as their own. Honest exploration of viewpoint and feeling, in oral and written form, was characteristic of the experimental students. Discussion of issues of friendship particularly elicited lively and frank debate. Should friends tell each other their most secret feelings and leave themselves vulnerable? Girls constantly advocated openness ("so that something can come of the relationship"); boys argued that being totally open was self-indulgent, led to dependency, and inhibited the development of autonomy. Such encounters were by no means unusual. The investigator felt confident that his students were growing in moral sensitivity and social understanding; he was less certain about whether a "vertical" measurement such as the MJJ, with its strong cognitive orientation, could capture what he perceived as the students' steady growth.

Methodological weaknesses of the study must be acknowledged. There are undeniable threats to internal and external validity in the design. Given the lack of randomization of students to groups, the absence of initial equivalency, the small sample size, the use of MJJ in written form, and the fact that the investigator taught only the experimental group, the results must be evaluated cautiously.

With respect to gains in the comparison group, unquestionably one of the strengths of the design of this study, as Higgins (1980) endorses in her review of moral interventions, is the duration of treatment—a full academic

year. Yet what constitutes a strength may also pose a threat to internal validity, for over such a time period it is possible for other factors to intrude. For example, DiStefano (1978), in her study of "The Psychology of Ethics in Relationships," achieved gains of 42 MMS in the experimental group and 26 MMS in the comparison group—similar to the results of this study. Because all her study's subjects were members of a democratic alternative program involving moral discussion and democratic decision making, it was hard to disentangle the effects of the intervention from the effects of the alternative program.

In this case, the teacher for the comparison group was a popular, lively, Socratic performer. The expectation was that the two classes would be equivalent, yet statistical analyses found them not to be so. It is possible, then, that there was a selection-maturation interaction; perhaps in light of their pretest inequivalency, the natural rate of change in the two groups would have been different even in the absence of the experimental treatment.

Directions for Future Research

Any further study undertaken along these lines should be conversant with the various threats, particularly to internal validity, that have been presented. The multifaceted nature of the treatment, with the design weaknesses already explained, render it hard to argue with certainty for the effects of the curriculum and impossible to make differential claims about the relative impact of parts within the curriculum. Yet, with the study by Colby et al. (1977), it offers tentative evidence that moral reasoning stage advance can be accomplished within an academic English literature course.

However, to be concerned only with the raising of MMS is too limited a goal. If moral development and moral education are more broadly defined, as Kohlberg (1979), Gilligan (1977, 1982), and Rest (1980) suggest, it must also cover moral sensitivity, moral character, and moral values and behavior. An English literature study, in that case, would wish to look at developmental variables other than just the Moral Maturity Score. Ego development may be a suitable variable. Loevinger (1976) has constructed a theory of ego development in which patterns of cognitive interpersonal and ethical development form successive and hierarchical world views: "In general, ego development is marked by a more differentiated perception of one's self, of the social world, and of the relations of one's feelings and thoughts to those of others" (Candee, 1974, p. 621). Fostering an understanding of concepts of self and the social world falls reasonably within the mandate of an English teacher concerned about moral growth.

In addition to developing other indices of change, a more rigorous study examining moral and ego development might look at the impact of different parts of a curriculum as manifested in writing samples throughout the year because connections between moral and ego stage and students' writing—critical, imaginative, reflective, and autobiographic—could help us to better understand the process of change and aid curriculum development.

Acknowledgments

The author acknowledges Guy Bramble, currently headmaster of Heritage Hall School, Oklahoma, with whom much of the curriculum discussed above was developed, and Carole Beal and Edward DeVos for their helpful comments on an earlier draft of this article.

References

- Blatt, M. (1970). *Studies on the effects of classroom discussions upon children's moral development*. Doctoral dissertation, University of Chicago.
- Blatt, M., & Kohlberg, L. (1975). The effects of classroom moral discussion upon children's level of moral judgment. *Journal of Moral Education*, 4(2), 129-61.
- Campbell, D., & Stanley, J. (1963). *Experimental and quasi-experimental designs for research* (pp. 47-50). Chicago: Rand McNally.
- Candee, D. (1974). Ego developmental aspects of new left ideology. *Journal of Personality and Social Psychology*, 30, 620-630.
- Colby, A., Gibbs, J., Kohlberg, L., Speicher-Dubin, B., & Candee, D. (1979). *Standard form scoring manual*. Cambridge, MA: Harvard University.
- Colby, A., Kohlberg, L., Fenton, E., Speicher-Dubin, B., & Lieberman, M. (1977). Secondary school moral discussion programs led by social studies teachers. *Journal of Moral Education*, 6(2), 90-111.
- DiStefano, A. (1978). *The psychology of ethics in relationships*. Unpublished doctoral dissertation, Boston University.
- Dozier, M. (1974). *The relative effectiveness of vicarious and experiential techniques on the development of moral judgment with groups of desegregated sixth grade pupils*. Unpublished doctoral dissertation, University of Miami.
- Garrod, A., & Bramble, G. (1977). Moral development and literature. *Theory Into Practice*, 16(2), 105-111.
- Gilligan, C. (1977). In a different voice: Women's conceptions of self and morality. *Harvard Educational Review*, 47, 481-517.
- Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge: Harvard University Press.
- Harris, D. (1976). *Psychological awareness and moral discourse: A curriculum sequence for moral development*. Unpublished doctoral dissertation, University of Wisconsin.
- Higgins, A. 1980. Research and measurement issues in moral education interventions. In R. Mosher (Ed.), *Moral education: A first generation of research and development* (pp. 92-107). New York: Praeger.
- Jennings, W., Kilkenny, R., & Kohlberg, L. (1983). Moral development theory and practice for youthful and adult offenders. In W. Laufer & J. Day (Eds.), *Personality theory, moral development and criminal behavior* (pp. 281-355). Lexington, MA: Lexington Books.
- Kohlberg, L. (1979). *Exploring the moral atmosphere of institutions: A bridge between moral judgment and moral action*. Heinz Werner Memorial Lecture presented at Clark University.
- Kohlberg, L. (1981). *Essays on moral development: Vol. 1, The philosophy of moral development*. San Francisco: Harper & Row.
- Kohlberg, L. (1984). *Essays on moral development: Vol. 2, The psychology of moral development*. San Francisco: Harper & Row.
- Ladenburg, T. (1977, March). Cognitive development and moral reasoning in the teaching of history. *History Teacher*, 184-198.
- Lieberman, M. (1979). *Teaching the Holocaust*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.
- Lockwood, A. (1978). The effects of values clarification and moral development curricula on school-age subjects: A critical review of recent research. *Review of Educational Research*, 48(3), 325-364.
- Loevinger, J. (1976). *Ego development*. San Francisco: Jossey-Bass.
- Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. New York: Basic Books.
- Rest, J. (1980). Basic issues in evaluating moral education programs. In L. Kuhmerker, M. Mentowski, & L. Erickson (Eds.), *Evaluating moral development* (pp. 1-12). New York: Character Research Press.

- Schaffer, P. (1974). *Moral judgment: A cognitive-developmental project in psychological education*. Unpublished doctoral dissertation, University of Minnesota.
- Scharf, P. (Ed.). (1978). *Readings in moral education*. Minneapolis: Winston Press.
- Sullivan, P. (1975). *A curriculum for stimulating moral reasoning and ego development in adolescents*. Unpublished doctoral dissertation, Boston University.

Appendix

A Sample of Curricular Material from the Huckleberry Finn Unit

Chapter 15 finds the companions separated in the fog and then reunited. The trick that Huck plays on Jim backfires and reveals a new dimension, a new depth, to the friendship. Huck says, “It was fifteen minutes before I could work myself up to go and humble myself to a nigger; but I done it, and I warn’t ever sorry for it afterward, neither.”

Suggestions for Exercises

- A. Invite students to explore Huck’s reasoning during the 15 minutes he spent making up his mind to apologize, then have them write a script, with imaginary dialogue, detailing the various arguments that culminate in the apology.

Suggestions for Study Questions

- A. Why is Huck so surprised by Jim’s indignant response to his innocent joke?
- B. Why does it take Huck a full 15 minutes of internal debate before he can bring himself to apologize?
- C. What considerations hold him back?
- D. As a friendship deepens, does the obligation to apologize for unkind or thoughtless acts or words become greater or lesser, or does it remain constant?
- E. Discuss the development of Huck’s role-taking ability by comparing the time when he does not apologize (the snakeskin incident in chapter 10) with the time that he does (after the separation in the fog).

Chapter 16 deals with the first of two major moral crises that confront Huck (the second arises in chapter 31). Huck’s dilemma involves the conflict between his friendship for Jim and his perceived obligations to society (and to Miss Watson in particular). His reasoning in resolving this dilemma is presented in the middle of the chapter and is a classic example of Stage 2 instrumentalism: doing “whichever comes handiest at the time.”

Suggestions for Study Questions

- A. What factors influence Huck’s initial decision to turn Jim over to the authorities?
- B. What considerations encourage him to change his mind when he decides to lie to the bounty hunters?
- C. Isn’t it wrong to lie to them? Is it wrong to lie under all circumstances?

- D. What are the moral implications of the rule Huck embraces when he decides in the future he will do “whichever comes handiest at the time”? How does this rule solve the problem of right and wrong for Huck?

Chapter 31 contains the moral climax of the novel. The comparison of Huck’s reasoning in chapters 16 and 31 reveals Huck’s moral development: Faced with essentially the same dilemmas, his moral reasoning moves from Stage 2 to Stage 3 as a result of the constant cognitive conflict besetting him. As expected, Huck’s moral reasoning vacillated between Stage 2 and Stage 3, but the dilemma posed by Jim’s flight has at last been resolved by a more adequate moral rule. The Stage 2 instrumentalism which characterized Huck’s reasoning in chapter 16 has now given way to a broader and less self-centered perspective: Huck’s new principle might be termed “friendship first.” The morally right is now defined as that which is compatible with friendship, and this broader social perspective has been made possible by his developing ability to take the role of others—to differentiate the *self* from the *needs* of self. Once he can make this differentiation, Huck is capable of distinguishing between others’ needs and his own.

Suggestion for Exercises

- A. Write the dialogue for a scene in which Huck and an imaginary character argue the morality of his decision to tear up the letter and set Jim free.

Suggestions for Study Questions

- A. Should Huck have torn up the letter?
B. Can his reasons for tearing it up be derived from the pragmatic rule formulated in chapter 16: doing “whichever comes handiest at the time”?
C. Does the principle underlying Huck’s decision to tear up the letter provide a more adequate basis for decision making than the principle of doing “whichever comes handiest at the time”?
D. Compare Huck’s commitment (loyalty) to Jim with their earlier relationship. What differences do you see, and how do you account for them?

ADELE THOMAS

Brock University

Problem Solving Among Pairs of Achieving and Learning Disabled Students

The present investigation of children's problem solving strategies and task perceptions sought to apply a processing model to examine individual differences in children's problem solving. Fifty-six pairs of 10-year-olds classified gifted, learning disabled or average achievers engaged in a river-crossing problem solving task using think-aloud procedures. Learning disabled (LD) students were further classified as either helpless or mastery oriented. Transcriptions of videotaped sessions were analyzed for use of heuristics and affective variables as well as for sequences of moves toward solution. While learning disabled pairs differed from gifted and average pairs in solution, frequency of recognition errors, bizarre moves, and aspects of strategy use, no differences were found between Gifted and Average groups. For the LD groups, deficits were identified in cognitive processes of comparison and evaluation, while metacognitive strategies for rereading or drawings indicated no differences compared with achieving students. No differences were found based on helplessness classification or for task-related affect and interaction. Implications for instruction pointed to the usefulness of pair procedures to assess problem solving among disabled learners.

In special education, interest in the problem solving abilities of learning disabled children has coincided with developments in cognitive psychology. Theories of information processing have contributed new perspectives on the role of metacognitive strategy deficits in such areas as problem identification and monitoring to explain the achievement of learning disabled children (Torgeson, 1986). At the same time, methodological advances related to the study of cognitive processes have been successfully applied to the study of learning disabilities. Protocol analysis and the use of problem solving pairs have begun to be successfully incorporated into research, assessment, and teaching practices for the learning disabled (Gelzheiser & Shepherd, 1986).

Much of the comparative problem solving research with achieving and learning disabled children has employed academic tasks (Stone & Michals, 1986). While this approach has provided insight into aspects of problem solving deficits for the learning disabled, it has also posed difficulties. It has not been possible to adequately match comparison groups on knowledge prerequisites when investigating academic problem solving. An alternative approach has employed a cognitive processing model based on tasks which require little academic background knowledge. It was the aim of the present study to investigate learning disabled children's use of problem solving

Adele Thomas is an associate professor in preservice and graduate studies at the College of Education with research interests in the area of learning disabilities.

strategies based on a model of cognitive processing. Before detailing this alternative, relevant research on the problem solving skills of learning disabled children is reviewed.

It would appear that disabled learners tend to disregard relevant information, do not adequately understand problems, do not use systematic planning, persist with unsuccessful alternatives, and do not evaluate their conclusions (Stone & Michals, 1986). Nevertheless, while the problem solving skills of learning disabled students have been studied over a variety of tasks from recall of written information and game rules to puzzle assembly, performance results have been somewhat mixed (Havertape & Kass, 1978).

Kavale (1980) sought to evaluate problem solving strategies on a reading comprehension task where students were to give reasons (aloud) for accepting or rejecting choices listed in a multiple choice format. While poor readers were three reading levels below that of average readers, there was no identification of the reading difficulty level of the material compared with average readers. Poor readers produced more comprehension errors, used reasoning strategies less frequently, and produced a higher rate of word recognition errors. It is difficult to ascertain the nature of the reasoning strategy deficit in a task environment where the deficit skill is integral to task completion.

From another perspective, Kotsonsis and Patterson (1980) sought to evaluate cognitive monitoring in learning to play a game. Each child was given one game rule at a time and was to signal when he or she thought he or she knew how to play. Learning disabled children, despite ability to recall game rules and attention to the task, judged themselves ready to play with significantly fewer rules available compared with average achievers. Although it was concluded that learning disabled children appeared less sensitive to the adequacy of information received and what was needed, there was no student verbalization to evaluate aspects of self-monitoring processes or student task perception.

Using academic and nonacademic tasks, Dweck (1975) and others (Butkowsky & Willows, 1980; Diener & Dweck, 1978) arbitrarily manipulated success and failure feedback to underachieving and achieving children. On receiving corrective feedback during problem solving, underachieving helpless children were characterized by greater use of disconfirmed hypotheses and ineffectual hypotheses compared with average achieving peers. The helpless children tended to discount their successes, were pessimistic about outcome, and made attributions for their difficulties to uncontrollable factors. Although dramatic differences have emerged in problem solving styles for the two groups, the authors have noted that these studies have included a strong evaluative component established by experimenter feedback of "right-wrong." Thus it remains unclear what and how problem solving processes were used effectively by able and disabled learners and to what extent on-going perceptions about one's performance or ability affect problem solving. Nevertheless, motivational factors and the syndrome of learned helplessness have been implicated as a significant influence in problem solving effectiveness of learning disabled students (Licht, 1984).

Harris (1982) addressed the issue of verbal self-monitoring and cognitive training as it affects the problem solving of young children. Learning disabled and normal achieving eight-year-olds attempted a puzzle task where one puzzle piece was rigged to fit incorrectly. LD students scored significantly lower in time on task than normally achieving children, while no differences were noted in the number of problem solving strategies used by both groups. At the same time, the proportion of task irrelevant private speech was found to be significantly higher for learning disabled than for achieving students. In addition, learning disabled students were more likely to use negative affective statements as part of the task monologue. Harris hypothesized that task persistence was related to task-irrelevant private speech, and subsequent cognitive training both increased use of task-relevant speech and persistence time. Nevertheless, the relationship between use of problem solving strategy and monitoring needs further clarification. In the Harris (1982) study, both groups of young children displayed similar use of a variety of observed strategies, and most of the children completed all of the puzzle that was possible.

It appears that cognitive monitoring has been a key area where problem solving differences have occurred. A consistent finding has been the use by learning disabled students of task irrelevant self-statements and negative self-evaluation in problem solving monitoring (Licht, 1984). Such negative evaluation tendencies might be expected for students who have histories of learning difficulties and who are characterized as helpless. When the problem solving task environment has an academic focus managed by an adult evaluator, the influence of negative evaluation may be heightened. However, in seeking to investigate the learning disabled child's capability to use cognitive processes for problem identification planning, and evaluation, it is necessary to provide a task setting which minimizes associations with school failure and adult intervention in the monitoring process. In this respect, the model developed by Simon (1975, 1983) and others (Jeffries, Polson, Razran, & Atwood, 1977) has particular merit in studying the problem solving of learning disabled students.

Problem solving is characterized by cognitive processes which operate in sequence of search activities (Simon, 1983). A solver searches a mental problem space sequentially, gradually adding to the store of information about the problem and its solution by considering essential features and using heuristics. Heuristics are tentative, general decision processes which organize and narrow the search for a problem solution. Means-end analysis is a prominent heuristic by which the solver compares the current state with the goal state, identifies differences, and readjusts or sets subgoals in order to reduce differences. Within means-end analysis the solver also sets up a priority system to decide which subgoals should be implemented. Memory processes are involved in storing information on those subgoals which have been met and those yet to be tried.

Simon (1975), in characterizing problem solving as a process of search, focused on the use of the well-defined problem. Well-defined problems contain in the problem statement all the information needed to determine what

the solution should be and to evaluate progress toward solution. In addition, well-defined problems usually require a standard sequence of problem decisions or moves. Processing models have been applied to a variety of well-defined problems including the Tower of Hanoi, water jug problems, and river-crossing problems, in which all information concerning restrictions and the goal state are specified (Jeffries et al., 1977; Klahr & Robinson, 1982). Because the use of a well-defined problem incorporates the cognitive processes of interest in studying problem solving and deemphasizes academic background, it would appear especially suited for an investigation of learning disabled children (Spitz, 1987).

The question of how effectively to record cognitive activity has been reviewed by Ericsson and Simon (1984). They have presented an overview of the validity of using verbal reports as data within the framework of information processing theory, and the use of concurrent, think-aloud reports has been widely accepted. Nevertheless, alternatives to think-aloud procedures have been suggested which maintain direct access to ongoing cognitive activity while allowing a more natural task environment for spontaneous verbalization. Problem solving in pairs and in small groups has been considered in light of their advantages for verbalization (Noddings, 1982; Whimbey & Lochhead, 1979). In particular, structuring the task for pairs seems well suited to eliciting decision making processes. Schoenfeld (1982) found that single-person situations were more apt to generate a set to produce something for the researcher, while two-person sessions eased pressure on subjects. When experimenter intervention was minimal, discussions between members of the pair regarding what to do next made the use of heuristics more explicit and observable so that the researcher was required to make fewer inferences about how decisions were made.

In summary this study sought to apply a cognitive processing model and solver pairs procedures to examine individual differences in children's problem solving strategies such as means-end analysis, problem understanding, use of task restriction rules, and use of subgoals. Because the study was concerned with motivational as well as cognitive aspects of problem solving, personal performance evaluations were examined in relation to effective use of problem solving variables, and an attempt was made to identify helpless student problem solving characteristics with a learning disabled sample.

Method

Overview

Three groups of average achieving, learning disabled, and gifted students were videotaped during a problem solving task conducted in pairs. Learning disabled students were designated either helpless or mastery-oriented based on questionnaire results and teacher ratings. At task completion, children were interviewed concerning their reasons for solution and perceptions of performance. Aspects of problem solving were coded and analyzed from written transcriptions of videotapes.

Subjects

112 students (90 boys, 22 girls), aged 10-0 to 11-6 years were selected from elementary schools serving predominantly suburban, middle-class neighborhoods in a large public school district in the Niagara region of southern Ontario.

Gifted

Initially 46 students were identified as gifted according to school admission criteria for enrichment programs. These students attended three enrichment classes which provided a minimum of half-time instruction in a segregated class setting. Students were categorized as gifted based on two of three subtest scores equal to or above 130 on the Canadian Cognitive Abilities Test. Teacher recommendation, student and parent inventories were also part of the school gifted identification process. Eleven male pairs and 12 female pairs were first identified based on student choice and teacher judgment of positive sociability for the pairs.

Because sex composition of the learning disabled population favors males, a ratio of approximately 4:1 was maintained for the gifted group by finally selecting all of the 11 male pairs and randomly selecting three of the female pairs from the initial pool of 46 students.

Learning Disabled

Fifty-six students were selected based on school board nomination for specific learning disability classes. Requirements were based on IQ scores equal to or above 85 and on academic achievement at least one year below grade level in one or more academic areas. All learning disabled students between 10-0 and 11-6 years of age (mean age, 10-4) were selected from seven full-time, segregated classes for children with specific learning disabilities. Twenty-two male and six female pairs were formed based on student choice, teacher estimate of sociability, and similar helplessness scores. A helpless LD group consisted of 13 male and three female pairs; a mastery oriented LD group consisted of nine male pairs and three female pairs.

Average Achievers

Twenty-eight students with a mean age of 10-6 and helplessness scores of 8 or above (Dweck, 1975) were selected at random from three grade 5 classes serving the same area from which gifted and learning disabled students were selected. Only those whose school records indicated achievement at grade level and no history of learning problems were selected for participation. The average achieving group consisted of 12 male and two female pairs based on student choice.

Measure of Learned Helplessness

Research on helplessness in learning disabled children has identified the significant role of attributions in perceptions of failure (Diener & Dweck, 1978; Dweck, 1975). Following assessment procedures developed in this research, the Intellectual Achievement Responsibility (IAR) Scale was administered to LD students two weeks prior to the experimental task

(Crandall, Katkovsky, & Crandall, 1965). Thirty-four items describe negative or success achievement experiences that occur frequently in children's daily lives. These are followed by attribution choices to either internal responsibility (ability, effort) or external, environmental causes. Ten items of the IAR refer to failure experiences in which effort is an alternative.

Based on the format developed by Dweck (1975) to identify helpless children, LD students were classified helpless who scored 7 and below on 10 effort items of the IAR, while those scoring 8 and above were designated mastery oriented. In the present sample, the mean effort score for LD helpless (LDH) pairs was 4.04 ($SD=2.12$) with scores ranging from 0 to 7. For mastery oriented (LDM) pairs the average effort score was 8.12 ($SD=.33$), with scores ranging from 8 to 9.

In addition a teacher rating scale adapted from Dweck (1975) was used to rate LD student helplessness. Fifteen items describe student behavior under failure and success situations on a five-point scale. Mean teacher ratings were 45.2 and 41.28 for LDH and LDM groups respectively, indicating similar helplessness ratings by teachers for both LD groups.

Task and Materials

The problem was a cover story for children of the traditional river-crossing problem where four travellers must cross from one side of a river to the other, under certain constraints (Fletcher, 1968). The problem was presented to all pairs by means of audiotape and in writing as follows:

Farmer Jones went to the County Fair with his prize goat, his pet wolf and his wife's prize pie. He had to cross a river in a small boat which fit only the farmer and *one* other object. If left together, the wolf would eat the goat, but not the pie. The goat would eat the pie if left alone with it. How should Farmer Jones work it out so that he can get all his possessions across the river?

Procedures

Three female experimenters, equally represented across the four student groups, conducted the sessions over a two-week period. Within each school, each pair was brought to a room which was set up with a table, chairs, cassette tape recorder for the children, and video equipment. It was explained that new school materials were being developed and it was necessary first to try them out with students to get their opinions.

Before attempting the experimental task, children practiced thinking aloud together with a training task involving two number completions. Children heard the problem on audiotape and followed along on a problem card. The experimenter explained that once started the children would be on their own. Students were encouraged to work together and talk aloud about the problems. They were reminded to use the audiotape and problem card as often as needed. The experimenter played the practice tape and modeled a practice problem using audio tape checking, self questioning, goal stating, and drawing on paper.

Experimental Task

Pilot testing revealed that 35 minutes was sufficient to allow students to generate alternatives or arrive at a solution. After 20 minutes, if no solution was achieved, one clue was given, "Remember, the farmer may take a possession with him when he goes *back* across the river." This clue represented the point at which students often have greatest difficulty, a move which is necessary but which, in the short term, is a move backward, away from solution (Simon, 1975). There was no interaction between experimenter and pair until the pair signaled they had a solution. If an incorrect solution was given, the experimenter offered only the feedback "no, that's not it."

Coding Procedures

Verbatim written transcriptions of videotaped sessions included all dialogue and self-talk within pairs. As well, nonverbal behavior such as listening to the tape recorder or using worksheets was noted on the transcriptions.

Coding Scheme

An adaptation of the Jeffries et al. (1977) model of the problem solving process formed the basis of the scheme used to analyze pair protocols. Each protocol was parsed into three types of unit. Move units constituted dialogue related to the selection of a move to one side of the river including evaluative discussion related to that move. State 0 units were those where students considered possible moves but were unable to select a move. State 8 units were those which summarized or reviewed after a solution was achieved. Total units constituted the sum of the three unit types. However, only State 0 and move units were considered for analysis in this study.

Move units were analyzed according to state and legality. According to Jeffries et al. (1977), a series of eight states are possible leading to the goal state of crossing the river. Each state represents one move. Moves at each state were *legal* if they conformed to the three constraints that neither goat and pie nor wolf and goat were to be left alone, and Farmer Jones was to cross with only one object. *Illegal* moves were those which violated any of the constraints. Because it was possible to continue to attempt a solution after making an illegal move, those subsequent moves were labeled *recognition errors*. The recognition error path led to premature solution at State 4. State 3 was a crucial move because it required something new, returning to the left bank *with* an object. The remaining states required recursion while adhering to the three constraints. Unit coding involved identifying the state of the unit and whether the move was legal, illegal, recognition, or bizarre error. The latter moves involved alternatives completely outside the constraints of the problem such as eating the pie or swimming across the river.

State 0 Coding

A coding scheme for problem solution processes at State 0 was developed for monitoring and heuristic activities as well as social aspects of problem solving in pairs.¹ Following Schoenfeld's (1982) study of these units as transition points in problem solving where difficulty may be encountered, it was of interest to identify types of strategy used as well as related affective and

social variables. Three problem solving strategies and two social components were identified which were rated according to level of response as follows:

Considering moves—ratings (1 to 4) based on the degree to which problem constraints were considered.

Problem identification—ratings (1 to 4) based on the degree of recognition of difficulty and specific action in seeking clarification of the problem.

Representing the problem—ratings (1 to 3) based on the degree to which the pair seeks to draw or to take notes about the problem.

Interaction style—ratings (1 to 5) based on the degree to which pair discussion is mutual/independent, helpful/unhelpful in generating ideas, listening, encouraging the other.

Affect—ratings (1 to 3) based on the degree to which the pair expresses enjoyment, interest, a sense of accomplishment.

Coder Reliability

Two coders who were unaware of experimental group classification received instruction in the coding scheme, in two phases, over a seven-day period. In phase one, using training transcriptions of pairs not included in the study, coders achieved 90% agreement in identification of types of units and moves. Coders proceeded to score transcriptions for only these phase one variables. Fourteen randomly selected protocols, representing 25% of the sample, were coded by both coders. Percentage of agreement between judges for the phase one variables ranged from .93 to .97 across the unit and move variables.

In phase two training, coders achieved a minimum of 90% agreement for ratings of protocols and for scoring of posttask performance evaluations. After training, phase two scoring was undertaken. For the previously randomly selected protocols, correlations between scorers ranged from .81 to .92 for the phase two variables.

Results

At the outset, groups differed significantly in solution of the problem ($X^2=30.77$, $df=21$, $p<.001$). While all Gifted pairs and most (11 of 14) Average pairs solved the problem, few of the LD pairs (4 LDM, 1 LDH) did so. There were few or no claims that solution was achieved through State 3 for Gifted and Average pairs. In contrast, LD groups made claims for solution at State 1 and State 2, with significance at State 2 ($X^2=18.32$, $df=9$, $p<.05$). At State 4, while all groups made some claims for solution, the LD groups made significantly more claims ($X^2=34.16$, $df=21$, $p<.05$). It will be recalled that State 4 represents two solution paths based on recognition errors. A closer analysis of move patterns will clarify the underlying basis for group differences in claims.

Move Types Across Problem States

Before analyzing group differences for each type of move across problem states, an initial MANOVA was performed to control overall alpha level within the study (Pedhazur, 1982). The Wilks' lambda approximate F value

was statistically significant for groups on all the move-related variables analyzed simultaneously, $F(12,130)=3.74, p<.001$.

Table 1 shows the group means standard deviations and F values for the four categories of move and indicates mean codable units produced in pair dialogues and mean frequency of problem solving impasse (State 0).

Significant group differences were noted for total units of dialogue produced. While LD groups did not differ from each other, they produced more total units than Gifted or Average pairs who did not differ. It would appear that while the LD pairs did not achieve solution, they engaged in more problem discussion than the two non-LD groups. However, no group differences were observed in instances of State 0 discussion (points of problem solving impasse) or frequency of legal moves.

Group differences were significant for all types of nonlegal moves (illegal, recognition, and bizarre errors). Table 2 summarizes mean moves by student group across six problem solving states.

Following the Jeffries et al. (1977) model, States 2, 6, 7 for illegal moves, States 1, 6, and 7 for recognition errors, and State 7 for legal and bizarre moves were not included in the analysis of moves across problem states. Either the type of move was not possible at that state, or there were too few moves made by groups. A groups by repeated measures (State) analysis was performed on frequency of each of the four move types in order to examine differences in group profiles as a function of problem state. Main effects due to problem state were significant for all types of move: F legal $(5,48)=29.94, p<.0001$; F illegal $(3,50)=9.77, p<.0001$; F recognition $(5,480)=8.79, p<.0001$; F bizarre $(4,49)=7.36, p<.0001$. Interaction effects for problem

Table 1
Group Means, Standard Deviations, and F Values
for Total Move-Related Variables

Variables	Gifted	LDH	LDM	Average	F Value
Units	13.07 (6.37)	25.92 (8.71)	23.44 (8.17)	19.07 (10.91)	5.52*
State 0	3.64 (3.59)	5.25 (3.28)	5.63 (3.20)	5.29 (5.37)	.72
Legal Moves	7.50 (1.74)	6.08 (3.70)	7.00 (4.10)	9.64 (3.91)	2.48
Illegal Moves	.21 (.58)	3.17 (2.59)	1.75 (1.77)	1.29 (1.73)	5.53*
Recognition Errors	.64 (1.65)	4.58 (3.12)	4.06 (3.57)	1.79 (3.17)	5.35*
Bizarre Moves	1.14 (1.79)	7.33 (6.20)	5.00 (4.87)	1.14 (1.35)	6.62**

* $p<.005$

** $p<.001$

Table 2
Group Means and Standard Deviations
for Type of Move Across Problem States

Move Type	State 1	State 2	State 3	State 4	State 5	State 6
<i>Gifted</i>						
Legal	2.29 (.91)	1.43 (.85)	1.21 (.42)	1.00 (.39)	.64 (.49)	.93 (.27)
Illegal	.14 (.53)	-- --	.00 (.00)	.07 (.27)	.00 (.00)	-- --
Recognition Error	-- --	.44 (.53)	.14 (.36)	.36 (.84)	.00 (.08)	-- --
Bizarre	.57 (.76)	.36 (.63)	.07 (.27)	.14 (.36)	.00 (.00)	.00 (.00)
<i>LD Helpless</i>						
Legal	3.42 (1.83)	1.67 (1.56)	.42 (.79)	.25 (.45)	.17 (.39)	.17 (.39)
Illegal	1.92 (1.98)	-- --	.25 (.62)	.67 (.78)	.17 (.39)	-- --
Recognition Error	-- --	1.75 (1.66)	.20 (.40)	1.95 (1.31)	.45 (.90)	-- --
Bizarre	2.75 (2.96)	2.42 (1.73)	.16 (.58)	1.42 (1.68)	.17 (.39)	.25 (.45)
<i>LD Mastery</i>						
Legal	3.94 (1.98)	2.69 (1.77)	.13 (.34)	.13 (.34)	.00 (.00)	.13 (.34)
Illegal	1.31 (1.50)	-- --	.13 (.34)	.31 (.60)	.00 (.00)	-- --
Recognition Error	-- --	.94 (1.18)	.94 (1.48)	2.00 (1.93)	.13 (.34)	-- --
Bizarre	1.31 (1.81)	1.38 (1.89)	.31 (1.01)	.88 (.96)	.38 (.81)	.50 (.89)
<i>Average</i>						
Legal	4.29 (2.20)	2.43 (1.45)	.79 (.43)	.79 (.43)	.71 (.46)	.64 (.50)
Illegal	.57 (1.28)	-- --	.36 (.75)	.36 (.63)	.00 (.00)	-- --
Recognition Error	-- --	.50 (1.09)	.50 (1.16)	.64 (1.15)	.14 (.36)	.36 --
Bizarre	.79 (1.12)	.21 (.42)	.07 (.27)	.07 (.27)	.00 (.00)	.00 (.00)

state by group were significant for F legal (15,133)=2.60, $p<.002$ and F recognition (15,133)=2.14, $p<.01$.

Based on simple effects analyses (Table 3), there was a tendency for Average and LD pairs to select more legal moves at States 1 and 2 than did Gifted pairs. From States 3 to 7 the trend was for Average and Gifted groups to produce more legal moves than LD groups who did not differ from each other.

For illegal moves, State 1 was the only state in which significant differences were noted, with both LD groups producing more moves than Gifted or Average pairs. States 1, 2, and 4 produced significantly more bizarre responses for LDH pairs than for either Average or Gifted pairs. For recognition errors, LDH pairs produced significantly more such moves than Average or Gifted groups at State 2. At State 4, both LD pairs offered more recognition errors than Gifted pairs, with LDM pairs also producing more than Average pairs. In Figure 1 the group profile for recognition errors is representative of the pattern for nonlegal moves (illegal and bizarre) across problem states. In the early States 1 and 2, while both LD groups produced almost twice as many nonlegal moves as Average pairs, the Average pair responses were also twice as frequent as those of Gifted pairs. At State 4—a pivotal state which requires a shift in strategy—the nonlegal responses of both LD groups were more than twice those of non-LD groups. For legal moves the pattern was reversed. While all pair groups produced more legal than nonlegal moves at States 1 and 2, Gifted pairs made the least legal moves. On the other hand, for States 3 through 6, Average and Gifted pairs maintained legal move production while LD production of legal moves was negligible. It appeared that while Gifted and Average pairs became more selective in their choice of legal moves, LD pairs shifted to nonlegal alternatives or returned to an earlier problem state. This conclusion is supported by the significantly greater LD claims for solution at early problem states.

Table 3
F Value Results for Group Simple Effects Analysis of Move Data

State	Legal Moves	Illegal Moves	Recognition Errors	Bizarre Errors
1	3.62*	4.17*	—	3.72*
2	3.83*	—	4.54**	7.47***
3	12.62***	1.31	1.96	.49
4	16.04***	2.19	5.24**	5.88***
5	11.75***	2.72	1.64	2.19
6	14.87***	—	—	—

* $p<.05$
** $p<.006$
*** $p<.001$

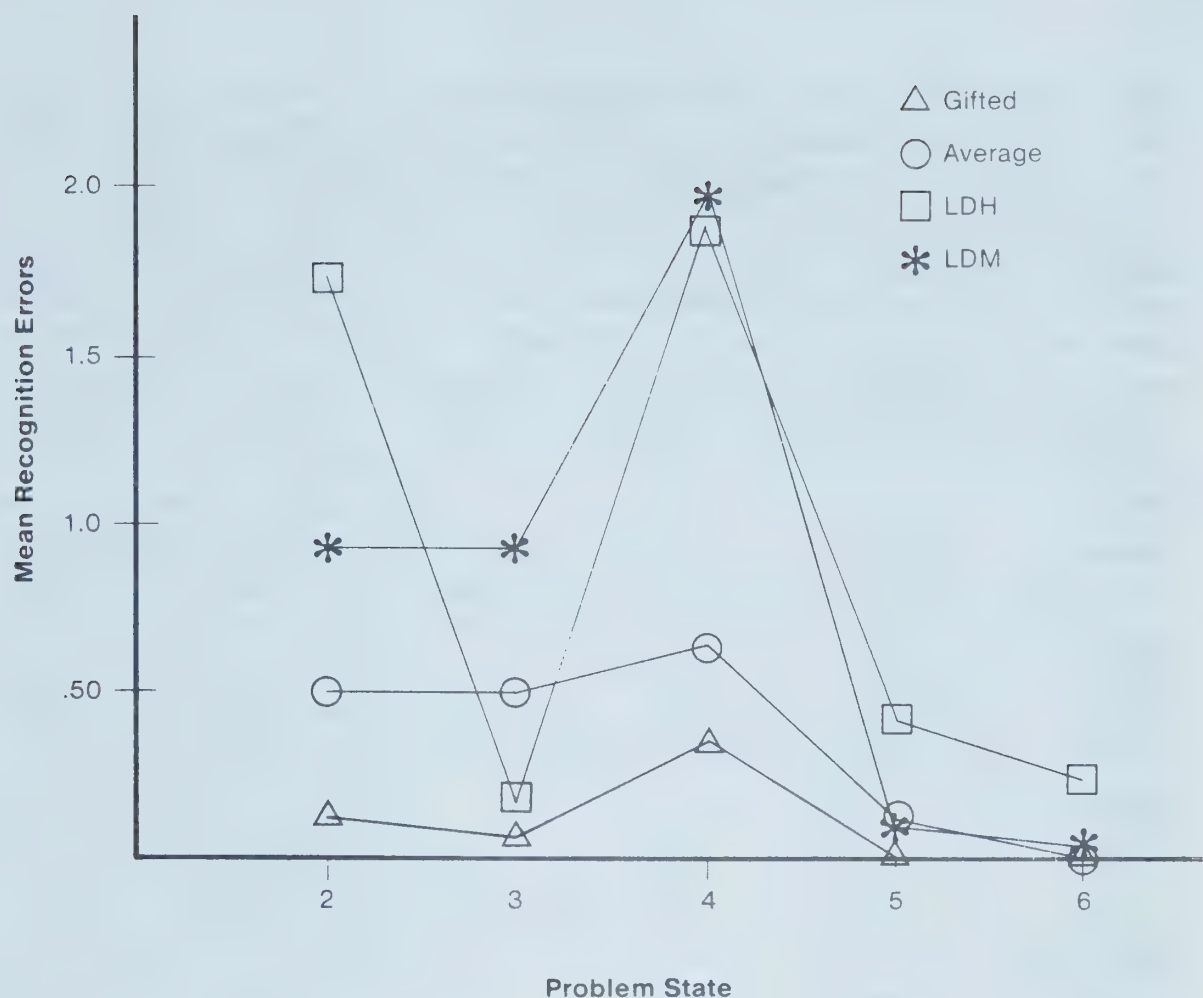


Figure 1: Graph of group recognition error means across states

Strategy Use and Interaction Style

Coded State 0 dialogue was analyzed by univariate tests on strategy and interaction style scores (Table 4).

Of the three heuristic strategies, Consider Alternatives was significant. While Average and Gifted pairs considered moves outside legal constraints, these pairs were more apt than their LD peers to reject such moves. On the other hand, LD pair dialogue was characterized by acceptance of nonlegal options.

For the Checks Problem and Representation (Draw) strategies, results indicated that LD and non-LD pairs used rereading, listening again to the audiotaped problem, and sketching the problem to the same extent. However, given the significant solution difficulty that LD pairs experienced in relation to non-LD pairs, this difference was not reflected in proportional differences in strategy use.

In considering affect and interaction scores across groups, pairs tended to be task-oriented and helpful in generating ideas and questions with each other. The mean group ratings for interaction indicate task involvement among pairs which was characterized by helpful dialogues and mutual problem solving, regardless of achievement label.

Table 4
Group Means, Standard Deviations, and F Values
for Strategy and Interaction Variables

Variable	Gifted	LDH	LDM	Average	F Value
<i>Consider</i>					
Mean	2.89	2.15	2.47	2.77	4.31*
SD	.67	.39	.40	.63	
<i>Check</i>					
Mean	1.42	1.80	1.64	1.55	.83
SD	.64	.53	.56	.59	
<i>Draw</i>					
Mean	1.31	1.27	1.38	1.14	.53
SD	.59	.36	.47	.21	
<i>Interact</i>					
Mean	4.94	4.95	4.91	4.88	.79
SD	.19	.08	.17	.25	
<i>Affect</i>					
Mean	1.72	1.64	1.71	1.56	.57
SD	.35	.22	.26	.39	

* $p < .009$

Discussion

The present study applied a processing model to the investigation of problem solving among children with different achievement backgrounds. Profiles of pair moves across problem states (Figure 1) indicated that earlier problem states were entered more frequently and were characterized by testing both legal and illegal alternatives. In later problem states solution was accomplished by a series of recursive steps which contained fewer legal possibilities. For all pairs, the last half of the problem space was characterized by a decline in the number of moves. All pairs, regardless of achievement classification, followed a one-move-at-a-time approach in the first three states by selecting a move and, only after selection, by performing some sort of evaluation test on their selection. Thus, with respect to the basic structure of the problem solving process, learning disabled pairs were not unlike gifted or average achieving peers.

Postmove evaluation determined subsequent move selection where recognition of past moves and consideration of problem constraints affected legality of choices. It was here that results of group differences pointed to critical distinctions between learning disabled and achieving solvers. Gifted pairs, in contrast to both Average and LD pairs used more efficient post-move evaluation and recognition of past moves while systematically check-

ing problem constraints. Learning disabled pairs, on the other hand, made more moves back to earlier states and included more illegal, bizarre, and recognition error moves than non-LD pairs. The higher rate of legal moves by LD and Average pairs than by Gifted pairs was a result of more frequent return to the start of the problem in order to try again. It may be concluded that both Average and LD solvers were less efficient than Gifted solvers, where efficiency is interpreted as selection of lowest number of legal moves for solution. However, Average solvers were also distinguished from their learning disabled peers in that LD evaluation processes were characterized by a significantly greater proportion of illegal, bizarre, and recognition errors at earlier problem states.

In observing Gifted and Average pairs, these groups engaged in frequent questioning and discussion which evaluated potential solutions against the goal state and against the three problem constraints. Their early use of this type of cognitive monitoring made unnecessary the later use of other metacognitive strategies such as rereading or sketching the problem. In the early stages of problem solving monitoring, LD pairs proceeded in a trial and error fashion. Although they used pair discussion and questioning to the same extent as other groups, LD pairs ignored constraints and made premature claims for solution based on unacceptable possibilities.

Thus it appeared that LD students were spontaneously able to initiate, at a minimal level, problem solving dialogue and common strategies. One strategy which proved more difficult for disabled learners was that of considering alternatives systematically. This strategy reflected ability to recall past moves rejected or accepted and to filter these for acceptability against the problem constraints. LD pair dialogue reflected less effective use of evaluation strategies and recognition processes than non-LD dialogue. Therefore, a distinction should be made between LD use of metacognitive strategies such as rereading, pictorial representation of the problem, or problem solving discussion and LD difficulty in more critical processes of comparing, recalling, and using common sense problem knowledge. LD solvers made significantly more attempts to reconstruct problem possibilities outside the given constraints. This finding points to a difference in what Simon (1983) has called commonsense problem knowledge based on an understanding that the problem statement is necessary and sufficient for solution. It appeared to be a significant factor in effective problem solving for disabled learners. While the greater incidence of bizarre and recognition errors in LD moves was indicative of difficulty in recalling constraints accurately, the concept of necessity to adhere to problem constraints also appeared to be less firmly understood by LD solvers than by non-LD solvers. Thus the present investigation underscored that even in a general search task, specific task knowledge and task-related cognitive processes were critical missing elements in LD problem solution.

Because of their greater apparent difficulty in using cognitive processes of comparison and evaluation to achieve a clear understanding of the problem, LD groups would be expected to rely on other compensatory strategies such as rereading and sketching the problem. Nevertheless, LD and non-LD

pairs used these strategies with similar frequency despite LD solvers' greater difficulty. It would appear that LD pair monitoring did not reflect adaptive or flexible use of "fix-up" metacognitive strategies, although these were minimally available in the repertoires of LD pairs. It would seem that the cognitive processing model used in this study was effective in more clearly delineating the nature of individual differences in problem solving. The relationship between cognitive and metacognitive processes in the problem solving of different groups of learners was further clarified based on the move-by-move requirements of the task presented and the spontaneous dialogue generated by the pairs procedure.

In the present study, helplessness was not a discriminating factor for LD groups, as no differential patterns emerged between the LD mastery and LD helpless groups. The finding of no differences on dimensions of negative affect and performance evaluation supported the hypothesis that a non-evaluative task environment which stressed cooperative problem solving minimized the negative influence of noncognitive factors on the problem solving of learning disabled students. Nevertheless, methodological limitations in the present study suggest caution in such conclusions. While helpless and mastery oriented LD students were identified based on the IAR, LD teacher ratings indicated similar judgments for both groups. Therefore, it is possible that both LD groups were similar on crucial aspects of helplessness despite their differential self-ratings. A recommendation for future research would include pretesting on challenge tasks to obtain behavioral indices of helplessness to supplement rating scales.

Implications for future research center on investigation of learner notions about problem solving and applications of the problem solving pair technique to other areas of problem solving. In much the same way that studies of metamemory have uncovered idiosyncratic understandings among disabled learners, relationships among notions about the problem solving process, failure experiences, and problem solving effectiveness await clarification. In a similar vein, the pair procedure as a way to generate task-related discussion and ideas in active communication is an area worthy of future research consideration. In the present study all solver groups seemed able to respond to the presence of a known peer with activities of listening, mutual sharing of ideas, and cooperative planning. From this perspective the problem solving activity and persistence levels of LD students were considerable, compared with the typical view of the LD learner as passive and dependent on teacher direction. It seems important to incorporate the potential inherent in cooperative peer arrangements in order to more accurately assess the problem solving capabilities of students and reflect the real-world setting in which much actual problem solving occurs. More recently the importance of peer dialogue in developing metacognitive skills has been recognized (Gelzheiser & Shepherd, 1986) and training programs have begun to incorporate the peer component in order to increase program effectiveness.

Note

1. A complete description of coding procedures and scoring criteria is available from Adele Thomas, College of Education, Brock University, St. Catharines, Ontario, L2S 3A1.

References

- Butkowsky, I., & Willows, D. (1980). Cognitive-motivational characteristics of children varying in reading ability: Evidence for learned helplessness. *Journal of Educational Psychology, 72*, 408-422.
- Crandall, V.C., Katkovsky, W., & Crandall, A.J. (1965). Children's belief in their own control of reinforcements in intellectual academic achievement situations. *Child Development, 36*, 91-109.
- Diener, C., & Dweck, C. (1978). An analysis of learned helplessness: Continuous changes in performance, strategy, and achievement cognitions following failure. *Journal of Personality and Social Psychology, 36*, 451-462.
- Dweck, C. (1975). The role of expectations and attributions on the alleviation of learned helplessness. *Journal of Personality and Social Psychology, 31*, 674-685.
- Ericsson, K.A., & Simon, H. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
- Fletcher, H. (1968). *Putting on your thinking cap*. Toronto: Abelard-Schuman.
- Gelzheiser, L.M. & Shepherd, M.J. (Eds.). (1986). Competence and instruction: Contributions from cognitive psychology. *Exceptional Children, 53*(2), 103-144.
- Harris, K. (1982). *The effects of cognitive training on private speech and task performance during problem solving*. Paper presented at the annual meeting of the American Educational Research Association, New York City.
- Havertape, J., & Kass, C. (1978). Examination of problem solving in learning disabled adolescents through verbalized self-instruction. *Learning Disabilities Quarterly, 1*, 94-99.
- Jeffries, R., Polson, P., Razran, L., & Atwood, M. (1977). A process model for Missionaries—Cannibals and other river-crossing problems. *Cognitive Psychology, 9*, 412-440.
- Kavale, K. (1980). The reasoning abilities of normal and disabled readers on measures of reading comprehension. *Learning Disabilities Quarterly, 3*, 34-45.
- Klahr, D., & Robinson, M. (1982). Formal assessment of problem solving and planning processes in preschool children. *Cognitive Psychology, 13*, 113-148.
- Kotsonsis, M., & Patterson, C. (1980). Comprehension monitoring skills in LD children. *Developmental Psychology, 16*, 541-542.
- Licht, B. (1984). Cognitive-motivational factors that contribute to the achievement of learning disabled children. *Annual Review of Learning Disabilities, 2*, 119-126.
- Noddings, N. (1982). *The use of small group protocols in the analysis of children's arithmetical problem solving*. Paper presented at the annual conference of the American Educational Research Association, New York City.
- Pedhazur, E. (1982). *Multiple regression in behavioral research* (2nd ed.). New York: Holt, Rinehart and Winston.
- Schoenfeld, A. (1982). *On the analysis of two-person protocols*. Paper presented at the annual meeting of the American Educational Research Association, New York City.
- Simon, H. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology, 7*, 268-288.
- Simon, H. (1983). Search and reasoning in problem solving. *Artificial Intelligence, 21*, 7-29.
- Spitz, H.H. (1987). Problem solving processes in special populations. In J.G. Borkowski & J.D. Day (Eds.), *Cognition in special children*. New Jersey: Ablex Publishing, 153-193.
- Stone, A., & Michals, D. (1986). Problem solving skills in learning disabled children. In S. Ceci (Ed.), *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities* (pp. 291-315). New Jersey: Lawrence Erlbaum.
- Torgeson, J.K. (1986). Learning disabilities theory: Its current state and future prospects. *Journal of Learning Disabilities, 19*(7), 399-407.
- Whimbey, A., & Lochhead, J. (1979). *Problem solving and comprehension: A short course in analytic reasoning*. Philadelphia, PA: Franklin Institute Press.

JOHN B. WOULDZIA

Simon Fraser University

Relationships Between Student Perceptions and Teacher Organization in Secondary-Level Alternate School Programs

The present study investigated the relationship between program organizational features and student perceptions of their learning environments in secondary-level alternate school programs. One hundred and three students enrolled in seven secondary-level alternate school programs responded to the Learning Environment Inventory which measures student perceptions of their educational environments, while program teachers completed the Classroom Characteristics Rating Scale which measures: (a) differentiation of instructional tasks among students, (b) student autonomy, (c) grouping, and (d) evaluation practices. A number of significant relationships—both desirable and undesirable—among specific program organizational features and student perception variables were identified. Differences regarding teachers' practices across the four organizational and instructional areas were also noted. Implications of these findings for teachers, curriculum designers, and other educational personnel are discussed.

Alternate education at the secondary level in Canada has been characterized by much organizational and curricular variation. This variation has been most noticeable among alternate school programs that specialize in providing instruction for students with socioemotional problems (Csapo & Gittens, 1979). One of the primary goals of secondary-level alternate education has been to provide a learning environment where students who have experienced academic failure in the traditional school system may enjoy academic as well as social and emotional growth (Csapo & Stevens, 1982; Stevens, 1985). In order to facilitate growth in these areas, educators responsible for planning alternate school curricula in a particular school district frequently design a program they believe will best meet the unique needs of their specific student population. Typically, this practice results in little program consistency from one alternate school to the next with respect to such areas as student autonomy, organization of instruction, and evaluation procedures.

In recent years, researchers have begun to acknowledge the importance of organizational features in educational settings and their influence on teacher and student perceptions of school-related events (Johnson, 1987). To date, however, most studies involving student perceptions of their learning environments have tended to focus almost exclusively on perceptual data and have ignored the importance of measuring the organizational fea-

John B. Woudzia is a graduate faculty associate in the Faculty of Education. His research interests include student motivation and cognitive strategies in teaching and counseling.

tures of these environments (Chavez, 1984; Fraser, 1985). Rather, organizational differences among programs have largely been assumed. Moreover, studies concerning teacher organizational and instructional practices in various types of school programs have focused on shared perceptions of ability (Rosenholtz & Rosenholtz, 1981; Rosenholtz & Wilson, 1980), academic and social power (Marx, 1985; Rosenholtz, 1981/1982), and academic self-concept (Simpson, 1981), rather than on student perceptions of classroom climate. Two notable exceptions are studies by Stevens (1985) and Gold and Mann (1986). Stevens examined the relationship between alternate school student perceptions of school climate and the organizational features that distinguished their schools. Although Stevens found that student perceptions of their learning environments were related to their alternate school programs' organizational structure, the program classification procedure used did not address teacher organizational and instructional practices per se. In Stevens' (1985) study, program organizational structure was more broadly construed by encompassing school goals, class size, source of funding, and target student population variables. Recent studies examining the relationship between student perceptions of school rules and program organizational features in conventional settings have tended to classify schools in a similar fashion by using broader categories such as size of school, type of school board, and number of student clubs or organizations rather than according to teachers' specific organizational and instructional practices (Macdonell & Martin, 1986).

In a similar study, Gold and Mann (1986) examined student perceptions of flexibility of school rules, commitment to the role of the student, and attitude toward school in alternate and conventional school environments at three different intervals across a 14-month period. Results indicated that alternate school students perceived their school programs as possessing greater rule flexibility, were more committed to meeting specific school expectations, and held more positive attitudes toward school than did their conventional counterparts. However, the "attitude toward school" variable in this study consisted of questions addressing four general components: global expression of liking of school, positive feelings for teachers, participation in extracurricular activities, and one's mood at the end of a school day. Furthermore, the main focus of this study was the comparative effects of alternate and conventional school environments on student perceptions or attitudes toward school, rather than the relationship between these perceptions and the organizational and instructional features of different alternate schools.

The purpose of this study was to investigate the link between student perceptions of their learning environments and the type of alternate school program in which students were enrolled. Two questions guided this inquiry:

1. Do alternate school programs differ in terms of their organizational and instructional characteristics?

2. What is the relationship between alternate school student perceptions of their learning environments and the organizational and instructional features that characterize these environments?

Method

Subjects

Data were gathered from 103 students in seven secondary-level alternate school programs in the Upper, Middle, and Lower Fraser Valley regions of southwestern British Columbia. Each of the seven alternate school programs was located in a separate building and in a different part of the school district from the regular junior high school to which it was attached administratively. In total, 71 boys and 32 girls were involved in the study. Students ranged from grades 7 to 10 with two students in grade 7, 25 students in grade 8, 49 students in Grade 9, 20 students in grade 10, and seven students who did not specify a particular grade level. All participating students were required to have attended their particular alternate school program for at least eight months prior to the beginning of the study. Program organizational data were also gathered from teachers at each of the seven schools. Only secondary-level alternate school programs which offered an academic core curriculum consisting of English, mathematics, science, and social studies were selected for inclusion in the study. General information regarding school curricula was furnished by teacher verbal reports of their instructional practices, as well as through policy and procedures manuals describing each of the school programs.¹

Instruments

Student perception measures. Student perceptions of their alternate school climate were measured by the Learning Environment Inventory (LEI) (Fraser, Anderson, & Walberg, 1982) which consists of 15 perception variables: cohesiveness, diversity, formality, speed, material environment, friction, goal direction, favoritism, difficulty, apathy, democracy, cliqueness, satisfaction, disorganization, and competitiveness. Alpha coefficients for individuals range from .54 for diversity to .86 for goal direction, while test-retest reliability for individuals ranges from .43 to diversity to .73 for friction. Intraclass correlations for groups range from .31 for diversity to .92 for formality and disorganization. A number of excellent reviews regarding the link between student perceptions as measured by the LEI and various learning outcomes are provided by Chavez (1984) and Fraser (1985). Studies involving the use of the LEI subscales as criterion variables have also been undertaken to measure the effectiveness of curricular innovations (Anderson, Walberg, & Welch, 1969; Fraser, 1978, 1979; Levin, 1980; Tisher & Power, 1976, 1978; Welch & Walberg, 1972), differences in student perceptions across various grade levels and subject areas (Anderson, 1971; Welch, 1979) as well as across different types of school programs (Randhawa & Hunt, 1981; Randhawa & Michayluk, 1975; Sharan & Yaakobi, 1981).

Program organization measures. Information regarding alternate schoolteachers' organizational and instructional practices were gathered

through an instrument called the Classroom Characteristics Rating Scale (CCRS). This instrument is essentially a modified version of an interview/questionnaire developed originally by Rosenholtz and Rosenholtz (1981) to measure dimensionality characteristics of classrooms across four areas: differentiation of task structure, student autonomy, grouping practices, and student evaluation. Teachers' organizational and instructional practices are viewed as either unidimensional or multidimensional in nature. Students in unidimensional classrooms are (a) instructed to use similar materials across subjects, (b) grouped collectively for the purposes of instruction, (c) rarely permitted to make decisions regarding type or length of activity, and (d) frequently evaluated on the basis of comparing one student's progress to that of another. Students in multidimensional classrooms, on the other hand, (a) use different materials across subject areas, (b) receive individual instruction, (c) are permitted a greater range of instructional choices or options, and (d) are evaluated on the basis of individual progress.

For the purposes of this study, the interview portion of the instrument was omitted, and the questionnaire portion modified to suit alternate school programs. The modification consisted of inserting the word "alternate" before the word "classroom" for the appropriate items. Teachers simply responded by checking *never*, *seldom*, *occasionally*, *frequently*, or *always* to items asking about their instructional practices across the four subject areas of English, mathematics, science, and social studies. Currently, no normative data or data regarding reliability and validity are available for either form of the CCRS. However, a number of studies have used the original instrument to explore the relationship between classroom organization characteristics and shared perceptions of ability (Rosenholtz & Rosenholtz, 1981; Rosenholtz & Wilson, 1980), academic and social power (Marx, 1985; Rosenholtz, 1981/1982), academic self-concept (Simpson, 1981), and attitude toward school (Rosenholtz & Simpson, 1984). While prior use of the instrument as evidenced by these studies illustrates the credibility of the instrument, as well as its ability to discriminate among organizational features of classrooms, the absence of data regarding validity or reliability is acknowledged as problematic with respect to the results reported in this study.

Procedure

Students completed the LEI in small group testing sessions in June. One morning was required to complete the data collection for four schools, while the remaining three schools required two consecutive mornings due to absenteeism. Groups consisting of a maximum of five students each were administered the inventory at each school. In a school where 12 students were enrolled, three groups of four students were tested. Information regarding student sex and grade was gathered. No other identifying information was requested.

Teachers completed the CCRS individually in June. Only one score for each of the four dimensions per subject area was required from each alternate school program. In schools where two or more teachers were employed,

consensus had to be reached concerning the instructional practices for each dimension before a single response for each item was recorded.

Results

Teacher Organizational and Instructional Differences

Results of the CCRS are presented in Table 1. Theoretically, the lower an individual school's score, the higher its degree of unidimensionality for the organizational dimension under consideration. Conversely, as the individual program scores increase across dimensions, the degree of multidimensionality also increases. Since it was possible to earn a total of 5 points across six subject areas for each dimension, a maximum of 30 points and a minimum of 6 points for each of the four dimensions was possible. The table indicates how each of the seven schools compared with respect to the four program dimensions.

Task structure. The program scores are uniformly low and quite similar along this dimension. The difference between the most unidimensional and multidimensional schools is 7 points (Program One with 9 points and Programs Three and Six with 16 points each). For this dimension, teachers were asked to respond to questions concerning how often their students use the same materials, such as texts, workbooks, and teacher-made materials across different subject areas. In general, instructional practices of teachers representing each of the alternate programs appear to be similar with regard to task differentiation. Furthermore, these practices tend to be unidimensional.

Student autonomy. Unlike task structure, there was much variability among program scores regarding student autonomy or decision making power. For this dimension, teachers were asked how frequently their students were permitted to make their own decisions regarding type and

Table 1
Results of the Classroom Characteristics Rating Scale

Programs N = 7	Program Organization Characteristics			
	Task Structure	Student Autonomy	Instructional Organization	Teacher Evaluation
1	9	8	10	18
2	11	10	11	22
3	12	15	16	30
4	13	17	21	24
5	16	20	17	26
6	16	18	25	30
7	12	25	28	29

Possible range of scores for each dimension is 6-30. Lower scores indicate unidimensionality; higher scores indicate multidimensionality.

length of activity for each subject. Teacher responses along this dimension ranged from *never* (two programs) to *frequently* (one program), with four of the seven teachers reporting that they *occasionally* permit their students to make these types of decisions.

Student grouping. Similar to autonomy, there was much variability among teacher grouping practices. For this dimension, teachers were asked how often they group the entire class together for instructional purposes. Again, individual school responses ranged from *never* (one program) to *frequently* (two programs) along the unidimensional/multidimensional continuum.

Teacher evaluation. Relative to the preceding dimensions, scores for all seven programs regarding teacher evaluation were uniformly high. For this dimension, teachers were asked how frequently they made comparative evaluations when assessing individual student progress across subject areas. None of the responses indicated a strong unidimensional preference. Teacher responses ranged from *occasionally* (one program) to *never* (three programs). The remaining three schools indicated that they *seldom* made comparative student evaluations.

Program Organization and Student Perceptions

The relationships among student perceptions and program organization data were explored to determine the extent to which teacher organizational and instructional practices relate to student perceptions of their alternate school programs. Program means and standard deviations for each of the student perception variables are presented in Table 2. Pearson correlation coefficients were computed for each of the four program organizational variables including a total value across each of the 15 climate variables. The program raw score on the respective organizational variable was correlated with the 15 mean student perception scores. A number of statistically significant correlations emerged from these analyses and are illustrated in Table 3.

Task structure. As mentioned, teachers tended to be unidimensional in this area by generally using similar materials for different students across subject areas. However, as these teacher practices become more multidimensional, significant correlations indicate that students like their material environment less ($-.74$) and perceive their programs as less formal ($-.86$), less goal-directed ($-.67$), and more disorganized ($.81$).

Autonomy. Differences regarding the amount of autonomy or decision making power program teachers typically give their students has been noted. As student decision making power increased, student perceived levels of formality ($-.81$), speed ($-.85$), difficulty ($-.81$), competitiveness ($-.93$), cohesiveness ($-.68$), diversity ($-.67$), and goal direction ($-.68$) decreased to a significant level, while perceived level of democracy ($.91$) increased significantly.

Grouping practices. Teacher grouping practices were widely dispersed along the unidimensional/multidimensional continuum. Results indicate that as instruction became more individualized, student perceived levels of

Table 2
Program Means and Standard Deviations for the
15 Student Perception Variables of the Learning Environment Inventory

Perception Variable		Program						
		1	2	3	4	5	6	7
Cohesiveness	$\bar{X}$	21.73	23.92	22.20	22.77	20.85	20.50	20.39
	SD	2.15	2.61	3.08	2.20	2.64	2.95	3.26
Diversity	$\bar{X}$	21.33	21.92	20.70	21.31	20.80	21.20	20.78
	SD	2.61	2.02	1.56	2.21	2.04	1.20	2.93
Formality	$\bar{X}$	25.33	25.25	21.00	21.00	19.73	19.60	21.04
	SD	2.02	1.82	1.94	2.51	2.75	2.01	2.92
Speed	$\bar{X}$	18.40	17.08	16.20	13.62	15.10	15.35	14.15
	SD	4.08	2.46	1.93	2.63	2.75	4.27	3.48
Material Environment	$\bar{X}$	21.60	23.08	21.90	20.92	20.55	19.40	20.96
	SD	1.92	2.11	1.37	3.28	2.68	3.24	2.79
Friction	$\bar{X}$	20.73	19.33	19.10	20.23	20.35	22.00	21.74
	SD	2.43	3.14	2.33	2.59	2.46	2.40	4.41
Goal Direction	$\bar{X}$	23.10	24.33	19.80	19.31	20.50	19.10	20.48
	SD	3.43	2.15	2.49	2.75	2.28	3.35	2.39
Favoritism	$\bar{X}$	16.10	12.67	16.40	15.08	18.33	17.70	15.96
	SD	3.49	3.06	2.59	3.20	3.16	2.87	2.99
Difficulty	$\bar{X}$	20.93	19.50	16.10	16.15	17.53	17.10	15.96
	SD	1.28	1.73	2.33	2.73	3.12	3.14	2.47
Apathy	$\bar{X}$	16.93	12.42	15.90	16.62	16.40	21.00	17.76
	SD	2.60	3.03	3.28	3.78	3.17	2.62	3.73
Democracy	$\bar{X}$	12.80	14.83	15.90	17.69	16.05	16.50	19.83
	SD	1.37	1.53	2.88	1.75	2.26	2.55	3.33
Cliqueness	$\bar{X}$	19.40	18.17	18.80	18.08	20.00	20.40	19.00
	SD	2.47	1.95	2.53	1.85	3.13	3.20	2.63
Satisfaction	$\bar{X}$	17.40	19.83	18.70	18.54	18.03	15.00	17.70
	SD	2.13	2.62	2.21	3.13	1.82	2.00	3.42
Disorganization	$\bar{X}$	12.87	11.92	16.00	14.38	16.65	16.65	15.09
	SD	3.27	2.23	3.13	2.43	4.15	2.56	3.50
Competition	$\bar{X}$	21.27	20.75	15.80	15.00	16.20	15.60	13.04
	SD	2.02	1.54	3.16	2.97	3.78	2.99	2.01

friction (.70), apathy (.68), and democracy (.90) increased significantly, while levels of perceived formality (−.75), speed (−.82), difficulty (−.79), competitiveness (−.91), attractiveness of the material environment (−.70), and goal direction (−.75) decreased significantly.

Student evaluations. Scores along this variable were more uniformly multidimensional than they were for the other three program organization variables. As multidimensionality increased, levels of student perceived for-

Table 3
Correlations Between Student Perceptions of their
Alternate School Programs and Teacher Organizational
and Instructional Practices (N = 7)

Student Perceptions	Program Organization Areas				
	Task Structures	Student Autonomy	Student Grouping	Teacher Evaluations	Total
Cohesiveness	-.49	-.68*	-.65	-.50	-.67*
Diversity	-.35	-.67*	-.48	-.64	-.62
Formality	-.86**	-.81**	-.75*	-.83*	-.89**
Speed	-.60	-.85**	-.82**	-.59	-.83**
Material Env.	-.74*	-.59	-.70*	-.49	-.70*
Friction	.33	.52	.70*	.25	.55
Goal Direction	-.67*	-.68*	-.75*	-.75*	-.80**
Favoritism	.64	.47	.38	.46	.51
Difficulty	-.53	-.81**	-.79*	-.85**	-.86**
Apathy	.50	.47	.68*	.47	.61
Democracy	.40	.91**	.90**	.68*	.88**
Cliqueness	.57	.25	.25	.30	.34
Satisfaction	-.47	-.28	-.52	-.34	-.45
Disorganization	.81**	.69*	.61	.81**	.79*
Competitiveness	-.56	-.93**	-.91**	-.82**	-.94**

* $p < .05$, ** $p < .01$

ality (–.83), difficulty (–.85), competitiveness (–.82), and goal direction (–.75) decreased significantly, while perceived levels of disorganization (.81) and democracy (.68) increased significantly.

Total. When data from all four program organization variables were considered collectively, a number of significant relationships—both desirable and undesirable—emerged. Specifically, as programs became more multi-dimensional across all four organizational areas, students perceived their programs as less formal (–.89), less rushed (–.83), less difficult (–.86), less competitive (–.94), and more democratic (.88). Conversely, they also perceived their alternate school environments as less physically attractive (–.70), less cohesive (–.67), less goal-directed (–.80), and more disorganized (.79).

Discussion

The findings of the present study indicate that teachers, administrators, and other educational personnel responsible for planning alternate school curricula should review the organizational and instructional features of alternate school programs if they wish to create learning environments that are perceived positively by their students. As noted, students perceived their programs as informal, comfortably paced, democratic, not difficult, and noncompetitive as multidimensionality increased across all organizational areas. Less encouraging were results indicating that students also perceived their learning environments as uncohesive, physically unattractive, lacking goal direction, and disorganized. These findings suggest that multidimensional instructional practices among alternate school teachers have aided in creating learning environments that are perceived as fair in most respects by their students. However, by creating these types of environments, it appears that teachers have sacrificed other equally important characteristics of their alternate school programs such as clearly established program goals, order or organization, and feelings of belonging or affiliation as perceived by their students.

To illustrate, it was discovered that student level of perceived cohesiveness, diversity, and goal direction is negatively correlated with greater multidimensionality across the organizational variable of student autonomy, while level of perceived disorganization is positively correlated with this same variable. These findings suggest that satisfaction among alternate school students decreases as opportunities to make individual curricular decisions increase. In other words, greater student autonomy regarding the selection of particular subjects or tasks is not necessarily associated with positive student perceptions of their alternate school programs. This finding is surprising because curricular inflexibility within the conventional school has frequently been cited as a major reason among alternate school students for dissatisfaction with this environment as well as their inability to cope with its demands (Gardener & Farnsworth, 1977; Gold & Mann, 1986). More importantly, however, it suggests that a more unidimensional or teacher-centered approach to instructional decision making may be preferable to an individualistic approach, especially in terms of enhancing student perceptions of their alternate school programs. By taking responsibility for deciding which activities students must perform, and by setting constraints on performance (e.g., number of exercises to be completed, amount of time per activity, etc.), teachers may have some positive effect on the perceptions alternate school students hold about their school programs.

It is also interesting to note that student perceived levels of friction and apathy are positively correlated with greater multidimensionality across the variable of student grouping, while perceptions of goal direction are negatively correlated with this same variable. In other words, as student grouping practices become less collective and more individualized, student perceptions of friction and apathy in alternate school programs increase rather than decrease. Perhaps more importantly, student sense of confusion regarding the actual purposes or goals of their alternate school programs also

increases as instruction becomes more individualized. These findings tend to corroborate existing data concerning the effects of individualized instruction which provides only marginal support for its usefulness in promoting either prosocial gains or more favorable affective responses toward schooling (Good & Brophy, 1986; Peterson, 1979). These findings seem especially important for alternate school teachers. In our discussions with the teachers who participated in this study, we discovered that many of them provided opportunities for individualized instruction simply because they felt that the class-as-group instructional approach typically found in conventional schools might have been perceived as too competitive, threatening, and impersonal by their students. Moreover, given the small number of students in their school programs, many of these teachers saw individualized instruction as not only a preferable alternative, but a manageable one as well. However, while little direct research regarding the comparative effects of individualized versus group instruction in secondary-level alternate school programs currently exists, the results of the present study as well as those described by Good and Brophy (1986) suggest that a more unidimensional approach toward grouping students for the purposes of instruction by teachers seems warranted.

To summarize, the results of the present study point to the need for alternate school personnel to become more sensitive to the relationship that exists between teacher organizational and instructional practices and student attitudes or perceptions of their programs. If teachers hope to create alternate school programs where social climates are perceived as differentially more supportive by students than traditional school climates, then a more unidimensional approach to certain features of program organization is necessary. Further research aimed at identifying more specifically those organizational and instructional practices that are linked to positive perceptions of classroom climate is needed to enable program personnel to modify and improve existing alternate school curricula.

Note

1. Further information about each of the seven alternate school programs may be obtained from John Woudzia, Faculty of Education, Simon Fraser University, Burnaby, British Columbia, V5A 1S6.

References

- Anderson, G.J. (1971). Effects of course content and teacher sex on the social climate of learning. *American Educational Research Journal*, 8, 649-663.
- Anderson, G. J., Walberg, H.J., & Welch, W.W. (1969). Curriculum effects on the social climate of learning: A new representation of discriminant functions. *American Educational Research Journal*, 6, 315-327.
- Chavez, R.C. (1984). The use of high inference measures to study classroom climates: A review. *Review of Educational Research*, 54, 237-261.
- Csapo, M., & Gittens, W. (1979). Survey of rehabilitation programs in British Columbia. *British Columbia Journal of Special Education*, 5, 227-253.
- Csapo, M., & Stevens, A. (1982). Do rehabilitation programs rehabilitate?: Teasing out the evaluability components. *B.C. Journal of Special Education*, 6, 351-383.
- Fraser, B. (1978). Environmental factors affecting attitudes toward different sources of scientific information. *Journal of Research in Science Teaching*, 15, 491-497.

- Fraser, B. (1979). Evaluation of a science-based curriculum. In H.J. Walberg (Ed.), *Educational environments and effects: Evaluation, policy, and productivity* (pp. 218-234). Berkeley, CA: McCutchan.
- Fraser, B. (1985). *Two decades of research on perceptions of classroom environment*. Paper presented at the American Educational Research Association Annual Meeting, Chicago.
- Fraser, B.J., Anderson, G.J., & Walberg, H.J. (1982). *Assessment of learning environments: Manual for Learning Environment Inventory (LEI) and My Class Inventory (MCI)* (3rd ed.). Perth: Western Australian Institute of Technology.
- Gardner, J., & Farnsworth, B. (1977). Evaluation of an alternate high school. *Reading Improvement*, 14, 299-302.
- Gold, M., & Mann, D.W. (1986). *Expelled to a friendlier place: A study of effective alternative schools*. Ann Arbor: University of Michigan Press.
- Good, T.L., & Brophy, J.E. (1986). *Educational psychology: A realistic approach* (3rd ed.). New York: Longman.
- Johnson, N.A. (1987). The pervasive, persuasive power of perceptions. *Alberta Journal of Educational Research*, 33, 206-228.
- Levin, T. (1980). Classroom climate as criterion in evaluating individualized instruction in Israel. *Studies in Educational Evaluation*, 6, 291-292.
- Macdonell, A.J., & Martin, W.B. (1986). Student orientations to school rules. *Alberta Journal of Educational Research*, 32, 51-65.
- Marx, R.W. (1985) Classroom organization and perceptions of student academic and social status. In I.E. Housego & P.P. Grimmett (Eds.), *Teaching and teacher education: Generating and utilizing valid knowledge for professional socialization* (pp. 75-109). Vancouver: University of British Columbia.
- Peterson, P. (1979). Direct instruction reconsidered. In P. Peterson & H. Walberg (Eds.), *Research on teaching: Concepts, findings, and implications* (pp. 57-69). Berkeley, CA: McCutchan.
- Randhawa, B.S., & Hunt, D. (1981). *Social-psychological environments and cognitive achievement*. Paper presented at annual meeting of American Educational Research Association, Los Angeles.
- Randhawa, B.S., & Michayluk, J.O. (1975). Learning environments in rural and urban classrooms. *American Educational Research Journal*, 12, 265-285.
- Rosenholtz, S.J. (1981/1982). Organizational determinants of classroom social power. *Journal of Experimental Education*, 50, 83-87.
- Rosenholtz, S.J., & Rosenholtz, S.H. (1981). Classroom organization and the perception of ability. *Sociology of Education*, 54, 132-140.
- Rosenholtz, S.J., & Simpson, C. (1984). Classroom organization and student stratification. *Elementary School Journal*, 85, 21-37.
- Rosenholtz, S.J., & Wilson, B. (1980). The effect of classroom structure on shared perceptions of ability. *American Educational Research Journal*, 17, 75-82.
- Sharan, S., & Yaakobi, D. (1981). Classroom learning environment of city and kibbutz biology classrooms in Israel. *European Journal of Science Education*, 3, 321-328.
- Simpson, C. (1981). Classroom structure and the organization of ability. *Sociology of Education*, 54, 120-132.
- Stevens, M. (1985). Characteristics of alternate schools. *American Educational Research Journal*, 22, 135-148.
- Tisher, R.P., & Power, C.N. (1976). Variations between ASEP and conventional learning environments. *Australian Science Teachers Journal*, 22, 35-39.
- Tisher, R.P., & Power, C.N. (1978). The learning environment associated with an Australian curriculum innovation. *Journal of Curriculum Studies*, 10, 169-184.
- Welch, W.W. (1979). Curricular and longitudinal effects on learning environments. In H.J. Walberg (Ed.), *Educational environments and effects: Evaluation, policy, and productivity* (pp. 167-179). Berkeley, CA: McCutchan.
- Welch, W.W., & Walberg, H.J. (1972). A national experiment in curriculum evaluation. *American Educational Research Journal*, 9, 373-383.

BOOK REVIEW

EDUCATION, RESEARCH, INFORMATION SYSTEMS AND THE NORTH. By W. Peter Adams (Ed.). Ottawa: Association of Canadian Universities for Northern Studies, 1987, 362 pp.

This is an excellent resource for students, teachers, researchers, and others interested in the Canadian North. Developed largely from the proceedings of a meeting of the Association of Canadian Universities for Northern Studies (ACUNS) held in Yellowknife in April 1986, the compilation focuses on five major areas of northern interest: education and research, higher education and training, teacher education, distance education, and information systems. The work also includes an extensive section on the northern research, teaching, and service interests of staff of the Association's 35 member universities.

The editor is to be commended for compiling and sorting such a wealth of information. Given the work's overall editorial excellence, one hesitates to find fault with what has been done; however, there are two areas that deserve some attention. First, Adams might have taken time in the introduction to delineate the various regions that are generally classified as being part of the Canadian North. Hamelin's northern categories are referred to in one article, and Finkler provides a definition that reflects an administrative research focus, but the term North remains undefined throughout the book. This raises a second question: Why were there contributions from some northern areas and not from others? The territories receive considerable attention, and there are valuable discussions on programs in Labrador and Arctic Quebec, but little is included on the northern areas of the other provinces. It would have been helpful, for example, to have had reports on the activities of the Northern Alberta Development Council and its counterparts in Saskatchewan and Manitoba. That not much is said in the work about provincial northern programs suggests that ACUNS operates with greater ease in areas where federal interests and responsibilities are paramount. Provincial commentaries, particularly those from the West, give scant attention to their own northern areas and instead devote much of their comment to affirming a willingness to serve areas above the 60th parallel. There is clearly a need for more collaborative research. And while ACUNS has undoubtedly attempted to moderate competing jurisdictional interests to effect coordinated and long-term research strategies, it is unfortunate that issues relating to the state of federal-provincial cooperation in northern matters are not addressed anywhere in the work.

The section on education and research provides an overview of services available to northern secondary students as well as of a number of special research and training programs sponsored by such agencies as the Yukon Science Institute. There is a consensus throughout that the elementary and

middle school programs provided for aboriginal children in the Territories are ineffective. Implicit in the reports of such hiring agencies as the Departments of Public Works and Renewable Resources is the assumption that until aboriginal children attain matriculation or equivalent status, they will not be able to participate in the sophisticated technical and scientific activities commonly associated with most northern ventures. The frustration experienced in participating in a failing educational system is exemplified in Tourangeau's comments on the needs of Métis children; in Higgins' discussion on the educational activities of the Inuit Tapirisat; and in Duggan, Johnson, and Naidoo's article on secondary school programs in the Northwest Territories. Some agencies, such as the Northern Heritage Society in Yellowknife, offer training programs for older Native children. But because the numbers involved are small and the programs appear to be continually on the brink of financial collapse, their overall benefit is minimal. Of much greater importance than references to such ventures as the Heritage project is that nothing substantial is suggested in the education section on how the school systems might be improved. Surely more is required than general affirmations about the need for an efficient school system.

The research presentations largely reflect developmental or applied research operations. While valuable in themselves, they have little to do with formal educational programs at the pre-university level. Indeed, the only research paper that deals with in-school programs is Stairs and Annahatak's article on Primary Inuktitut programs in the section on teacher education. Educational and other forms of human science research are clearly not a priority for either the universities or the federal government. As Finkler points out, less than 30% of the budget of the federal government's northern scientific training program was devoted to human science research by the universities during the period 1980-1985. It is not clear how much of this allocation was in support of school based research, but the presentations and other comments suggest that this area receives very little funding.

The sections on higher education and teacher education are more promising both in terms of program focus and outcomes. The articles by Stapleton and Cleveland describe the development of the Northwest Territories' Arctic College. This institution, with its several campuses, is providing a good range of programs despite such problems as staffing, course development, the readiness of applicants, and the vastness of the area served. There are frequent calls in the papers for interinstitutional cooperation. And while the matter of institutional autonomy is mentioned as a deterrent in establishing "links between Arctic College and the universities and colleges of the South" (p. 89), an equally important but unstated concern has to do with how the cost of such cooperation will be met.

Arctic College programs are based to a significant extent on the success of several earlier territorial postschool programs—the cooperative and teacher training initiatives especially. Lyall's article on Arctic cooperatives is somewhat disappointing because it ignores the role that missionaries played in promoting and supporting the movement and because it stresses

the general administrative rather than the product and delivery arrangements adopted by local communities. The accounts of the origins and growth of teacher training programs, on the other hand, are more comprehensive. One might fault Cram and Wilman for being overly enthusiastic about the accomplishments of the Kativik and EATEP teacher training programs, but there is little doubt that these programs, together with the initial teacher education project in Yellowknife, are excellent models for all areas of northern higher education.

Much current thinking about northern teacher education programs is based on the conviction that instruction should be carried out insofar as possible on a face-to-face basis at the community level. Although this delivery mode differs somewhat from those discussed in the articles on distance education, none of the approaches proposed involves students leaving their communities for extensive periods. Apart from Owen who provides an interesting profile of a typical distance delivery learner, the section's contributors do not have much to say about their client's ethnicity or social economic status. The articles by Hickman and Smith are particularly noteworthy in that they, unlike the other commentators, at least suggest using Native languages and local cultural orientations in determining course content.

The section on information systems demonstrates the vast amount of data that is available on the North in libraries, archives, and other sources, and the recently developed on-line systems that make it possible to have almost instantaneous access to much of this information. The section is troublesome nonetheless because it shows that one requires a fairly high level of both computer competence and funding support to use these systems effectively. This suggests that Natives and other northerners will find it increasingly difficult to obtain data on their own and that they will become more dependent on specialists not only for information but also for the interpretative frameworks in which such information will be used.

As stated earlier *Education, Research, Information Systems and the North* is an extremely valuable resource. From an historical perspective its contents indicate that a new form of northern development is in place, one that is much more comprehensive and far-reaching than such previous systems as the turn-of-the-century fur trade or the coming of government in the 1950s. The volume depicts an extensive range of knowledge and research systems and, in keeping with the editor's aspirations, it should be looked on as a basic reference for anyone interested in the North and especially by those who live and work in all parts of northern Canada. One can only hope that this source book will be the first of many such ACUNS publications.

Robert Carney
University of Alberta

For Contributors

Please prepare manuscripts in accordance with these guidelines.

- Format**
- Manuscript in triplicate.
 - Typewritten, double-spaced throughout.
 - Abstract of approximately 100 words, on a separate sheet.
 - Manuscript not to exceed 6,500 words excluding graphics.
 - Author's name and affiliation on a separate title page; title only on first page of manuscript to ensure anonymity in the review process.
- Referencing**
- Sources in parentheses after each reference (direct or otherwise), giving author's name (unless mentioned in text), year of publication, and page number(s).
 - Direct quotes of 40 words or less in double quotation marks incorporated into text; those longer than 40 words in indented block format.
 - All sources listed alphabetically at the end of the manuscript under the heading *References* using the APA style.
 - Explanatory notes, numbered consecutively and marked in the text with superscript numerals, may appear before the References under the heading *Notes*; any citations in notes follow the same format as other references.
- Graphics**
- Each table and figure on a separate page at the end of the manuscript.
 - Figures, charts, and diagrams in camera-ready form.
 - Number tables in Arabic numerals and center *Table 1* over each heading.
 - Number figures in Arabic numerals before the caption at foot of each graph, e.g., *Figure 1: ...*
 - Indicate placement of figures and tables in text, i.e., *Insert Table 3 here.*
- Style**
- *AJER's* editorial style conforms closely to the *Publication Manual of the American Psychological Association*.
 - Editorial changes may be made to manuscripts.
 - For spelling, consult *Webster's New Collegiate Dictionary*; spelling in quoted material remains as in the original.
- Diskettes**
- If manuscript is produced on an IBM, IBM compatible, or Macintosh computer, please also send diskette. Microsoft Word, MultiMate, WordPerfect, WordStar, Xerox Writer, XyWrite, and DisplayWrite 3 are acceptable.

Submission to *AJER* implies the manuscript is not being considered for publication elsewhere.

AJER is indexed in the *Alberta Education Index*, *Canadian Education Index*, *Current Contents/Social and Behavioral Sciences*, *ERIC/Current Index to Journals in Education*, and *Social Science Citation Index*; appropriate articles are abstracted in *Cognitive Development Abstracts*, *Educational Administration Abstracts*, *Educational Technology Abstracts*, *Family Abstracts*, *Language Behavior Abstracts*, *Multicultural Education Abstracts*, *Psychological Abstracts*, *Research into Higher Education Abstracts*, *School Organization and Management Abstracts*, and *Sociology of Education Abstracts*, and reviewed by the *Journal for Research in Mathematics Education*.

